7-20-2016

Image Quality Modeling and Optimization for Non-Conventional Aperture Imaging Systems

Philip S. Salvaggio
pss7119@rit.edu

Image Quality Modeling and Optimization for Non-Conventional Aperture Imaging Systems

by

Philip S. Salvaggio

B.S. Imaging Science, Rochester Institute of Technology, 2012
B.S. Computer Science, Rochester Institute of Technology, 2012

A dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy
in the Chester F. Carlson Center for Imaging Science
College of Science
Rochester Institute of Technology

July 20, 2016

Signature of the Author _____

Accepted by _____
Coordinator, Ph.D. Degree Program Date

CHESTER F. CARLSON CENTER FOR IMAGING SCIENCE
COLLEGE OF SCIENCE
ROCHESTER INSTITUTE OF TECHNOLOGY
ROCHESTER, NEW YORK

CERTIFICATE OF APPROVAL

Ph.D. DEGREE DISSERTATION

The Ph.D. Degree Dissertation of Philip S. Salvaggio
has been examined and approved by the
dissertation committee as satisfactory for the
dissertation required for the
Ph.D. degree in Imaging Science

Dr. John R. Schott, Dissertation Advisor

Dr. Robert D. Fiete

Dr. Jie Qiao

Dr. Leon Reznik

Date

Contents

List of Figures

Image Quality Modeling and Optimization for Non-Conventional Aperture Imaging Systems

by

Philip S. Salvaggio

Submitted to the
Chester F. Carlson Center for Imaging Science
in partial fulfillment of the requirements
for the Doctor of Philosophy Degree
at the Rochester Institute of Technology

Abstract

The majority of image quality studies have been performed on systems with conventional aperture functions. These systems have straightforward aperture designs and well-understood behavior. Image quality for these systems can be predicted by the General Image Quality Equation (GIQE). However, in order to continue pushing the boundaries of imaging, more control over the point spread function of an imaging system may be necessary. This requires modifications in the pupil plane of a system, causing a departure from the realm of most image quality studies. Examples include sparse apertures, synthetic apertures, coded apertures and phase elements. This work will focus on sparse aperture telescopes and the image quality issues associated with them, however, the methods presented will be applicable to other non-conventional aperture systems.

In this research, an approach for modeling the image quality of non-conventional aperture systems will be introduced. While the modeling approach is based in previous work, a novel validation study will be performed, which accounts for the effects of both broadband illumination and wavefront error. One of the key image quality challenges for sparse apertures is post-processing ringing artifacts. These artifacts have been observed in modeled data, but a validation study will be performed to observe them in measured data and to compare them to model predictions. Once validated, the modeling approach will be used to perform a small set of design studies for sparse aperture systems, including spectral bandpass selection and aperture layout optimization.

Chapter 1

Introduction

Predicting the image quality of a proposed or theoretical imaging system that uses an unconventional aperture is a difficult and unsolved problem. In the remote sensing community, this problem is of the utmost importance, as new imaging systems require both sizable monetary and time investments. These conditions are further exacerbated when considering space-based systems, such as WorldView 3, which had a 10-year $3.55 billion development budget [de Selding, 2010]. As remotely-sensed images were originally analyzed exclusively by humans, a task-based image quality metric was a viable approach. The National Image Interpretability Rating Scales (NIIRS) metric is one such system. In this system, an image may be ranked NIIRS 0 to NIIRS 9 depending on the analysis tasks that may be performed on that image. For instance, in order to obtain a NIIRS 6 rating on a visible image, an analyst must be able to identify a spare tire on a medium-sized truck [Leachtenauer, 1996]. While this system is viable, it requires a human to perform analysis tasks and thus introduces subjectivity. As computational resources have become abundant and remote sensing systems have made the switch to digital sensors, human involvement in image analysis has decelerated, while automated computational analysis has become increasingly popular. As such, keeping the human in the quality analysis loop for tasks now performed by computers is both inefficient and undesirable. In order to eliminate subjectivity and gain the ability to perform sophisticated design optimizations based on a quality metric, it would be desirable to have a predictive image quality metric that was based purely off of the design parameters of the system.

While the NIIRS system works well for predicting image quality in its use cases, it

is not a completely general image quality metric. Rather, image quality is determined by the ability of a remotely-sensed image to perform the task at hand. For instance, the majority of the NIIRS criteria involve the ability of images to provide military intelligence. However, if the task was instead land cover classification, the criteria for image quality would be much different. In this case, spectral resolution and signal-to-noise ratio would be more significant compared to ground sample distance. As these examples might suggest, the problem becomes increasingly difficult when the system under consideration is either multispectral or hyperspectral. While the issues with creating a perfectly generic image quality metric are likely insurmountable, NIIRS has shown that making a metric specifically tailored to a task is an effective tactic. This philosophy will be a driving factor behind the structure of the aperture design framework introduced in this research.

To perform trade space analyses when designing imaging systems, it is desirable to express the image quality metric in terms of the system's design parameters. In the case of NIIRS, an effort was made to do just this. A regression analysis was performed to express the anticipated NIIRS rating of images produced by a novel system in terms of parameters such as ground sample distance, edge response and signal-to-noise ratio [Leachtenauer et al., 1997]. This expression is known as the General Image Quality Equation (GIQE). The form of GIQE 4 is given as

$$\text{NIIRS} = 10.251 - a \cdot \log_{10}(GSD) + b \cdot \log_{10}(RER) + c \cdot H - d \cdot \frac{G}{SNR} \qquad (1.1)$$

where GSD is the ground sample distance, RER is the relative edge response, H is the edge overshoot due to sharpening, G is the gain of the sharpening filter, SNR is the signal-to-noise ratio of the image and a, b, c and d are linear weighting coefficients. If the GIQE is examined, it is clear that there are three factors that correlate with image quality: resolution, post-processing and noise. The a coefficient, which corresponds to the resolution term, has the largest weight, implying that resolution is the most important of these three factors.

Like the NIIRS system itself, the GIQE is a useful tool with some notable limitations. Due to its nature as a regression fit, there are inherent limits on its utility. Primarily, the fit is only valid over the domain of the input variables that was spanned by the input data set. As such, if new designs are created that push the "edges" of traditional optical

design, their resulting system parameters may fall outside the domain of the data set used to fit the GIQE. In this region, the fit may be unstable and the functional relationships may be incorrect. This limits the types of systems that can be analyzed. In addition to the domain of input variables, the choice of input variables also limits the utility of the GIQE. While the system parameters in the GIQE were sufficient to characterize the image quality of the systems used to fit the function, novel systems may have additional or different parameters that govern their performance. For instance, if the majority of the data set used to fit the GIQE were from imaging systems with monolithic apertures, then the GIQE might not be the best tool to predict the performance of sparse or segmented aperture systems. As will be examined extensively in this research, changing the shape of the aperture results in drastic changes in the appearance of the output image. Since the post-processing of these types of systems is more complicated, artifacts other than edge overshoot can arise in the output image, which further degrade image quality. So, while these systems can be characterized in terms of relative edge response, edge overshoot and filter gain, there is no guarantee that the functional relationship established by GIQE holds for non-conventional apertures.

As was previously mentioned, the GIQE indicates that the resolution of an imaging system is the dominant predictor of the ultimate quality of its images. While resolution in the GIQE is measured in terms of ground sample distance, the true nature of resolution is more complex and must be understood when analyzing systems with non-circular apertures. Under the Rayleigh resolution criterion, resolution is defined in terms of two identical point sources. As these point sources are brought closer together, there will come a point at which they will be indistinguishable from a brighter, single point source positioned directly between the two point sources. This situation is illustrated in Figure 1.1. The distance between these two point sources at which this confusion occurs is the resolution limit of that system. The Rayleigh criterion states that this resolution limit is directly determined by the size of the point spread function (PSF) of the imaging system, or the amount of blur it introduces [Goodman, 2005]. The size of the system's PSF can be determined by a large number of parameters; some of these parameters are controllable by the designer of the imaging system, while others are unavoidable.

In conventional optical imaging systems, the resolution of the image is theoretically

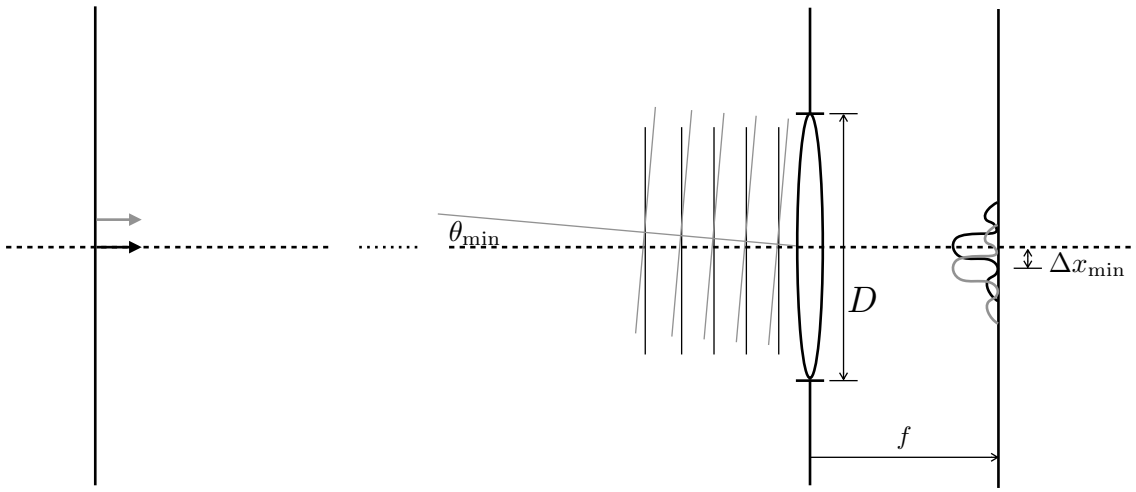


Figure 1.1: An illustration of the Rayleigh resolution criterion

limited by diffraction. Other factors, such as optical aberrations, detector sampling rates and atmospheric scattering, may also be limiting factors in imaging resolution. However, in the absence of all other factors, the conventional imaging system can never hope to achieve resolution greater than the diffraction limit for a single exposure. For an imaging system with a circular aperture function, the angular resolution limit of incoherent incident radiance due to diffraction is given by

$$\theta_{min} = 1.22\frac{\lambda}{D} \text{ [radians]} \tag{1.2}$$

where λ is the wavelength of the light being imaged and D is the diameter of the aperture of the imaging system. For applications that are imaging on a focal plane, this expression can be reformulated to give the resolution in terms of distance on the focal plane. This is done by using the focal length, f, of the system,

$$\Delta x_{min} = 1.22\frac{\lambda f}{D} = 1.22\lambda(F\#) \text{ } [m] \tag{1.3}$$

where $F\# = f/D$ is the F-number of the system. Examining these equations yields several conclusions. All things being equal, a system imaging at a shorter wavelength, should achieve better resolution. Similarly, a system with a larger aperture should have better res-

olution than a system with a smaller aperture. Given that resolution has been determined to be the dominant factor in image quality, it also stands to reason that the larger system will have better image quality, assuming the other terms in the GIQE can be held constant.

Practically speaking, holding focal length constant while increasing the diameter leads to a surface with increased thickness and sag. This often leads to an increase in the amount of optical aberration present in the optics, which can degrade image quality. In addition to an increase in aberration, larger apertures also come at the expense of the overall weight, volume and cost of the imaging system. For a traditional system, the cost tends to scale by at least the square of the diameter [AWMA and SPIE, 1996] While this is a problem for ground-based imaging systems, it is a prohibitive issue for space-based imaging systems, as launch vehicles can only take a given amount of mass and volume into space. Thus, until optics for space-based imaging systems can be fabricated and assembled in space, imaging systems will have to operate under a mass and volume constraint. This gives rise to the challenge of balancing the resolution of space-based systems against the mass and volume of the system. Ignoring the enclosures, electronics and other satellite components, this challenge is equivalent to maximizing the amount of resolution obtained from a given amount of glass or volume.

Achieving greater spatial resolution under mass and volume constraints is the motivation for the design of sparse aperture and synthetic aperture systems. The research described here will focus on the design and performance modeling of such systems. Performance modeling will occur in both the physical and computational domains. Computational modeling is desirable, as the performance of a large number of systems can be evaluated without constructing any optical components. However, in order to build trust in the computer model and base design decisions upon its output, validation experiments must be run. In this research, a small-scale sparse aperture system will be constructed in the laboratory for performing model validation. This system will be used to both validate intermediate outputs of the model, as well as artifacts that occur in the final imagery obtained from sparse aperture systems.

Chapter 2

Objectives

This chapter will outline the top-level objectives of this research and indicate the contributions of these objectives. As was mentioned in Chapter 1, this research will focus on the problem of sparse aperture image quality, specifically in a remote sensing situation. That is, the sparse aperture system will be focused at infinity and the input to the system will be polychromatic incoherent illumination. Given this imaging situation, the research proposed here will focus on two main efforts. The first of these is to perform a laboratory validation study of the sparse aperture system modeling methodology summarized in Chapter 4. Validation of monochromatic sparse aperture point spread functions have been attempted before, however, this research will aim to build upon previous studies. The main components of this research are:

- Design a laboratory optical system that can be used to simulate a sparse aperture system. This system should be able to:
  - Simulate a system with negligible wavefront error.
  - Introduce small amount of wavefront error in a characterizable manner
  - Provide controllable or characterizable broadband illumination to the imaging system.
  - Provide a mechanism by which to measure the system modulation transfer function.
  - Provide a mechanism by which extended scene analysis may be performed.
- Construct the designed laboratory system.

- Perform validation experiments on a sparse aperture system that can be replicated in the computer model.
  - Isolate the MTF contributions due to the aberrated pupil function.
  - Verify the predicted existence of post processing artifacts in observed imagery.

This validation study offers a number of advantages over existing studies. Previous studies have focused on measuring the monochromatic point spread function of a sparse aperture system. This study will also allow for the inclusion of spectral effects and wavefront error. This will allow a system designer to greater explore a larger portion of the sparse aperture design space using a validated model. The OTF validation experiment will be described in Chapter 5. In addition, this system will allow for the introduction of extended scenes into the system. Simulation of sparse aperture imagery in previous studies has indicated the possibility of unpleasant image artifacts due to post-processing. This research will allow for the verification of these artifacts in real imagery. The methods and results of this study will be presented in Chapter 6.

Once the modeling methodology has been validated, it can be used to predict the performance of theoretical sparse aperture systems. Thus, it can be a useful tool in trade studies involving sparse aperture system design. For example, if reconnaissance applications are being targeted, a ΔNIIRS study, like the one presented in [Garma, 2015], could be performed with the model's imagery predictions. Alternatively, if an automated analysis, such as crop coverage classification, is the targeted application, modeled imagery can be used as test data to assess the performance of a proposed telescope design. In this research, two design problems will be investigated: subaperture layout and spectral bandpass selection. These studies will both serve as demonstrations of how the validated model can be used to aid in design studies. However, because sparse aperture image quality is not yet a well-understood problem, these demonstrations are intended as starting points, and not as authoritative conclusions on optimal sparse aperture system design. The main components of this research are:

- Create a subaperture layout optimization algorithm.
- Validate the functionality of the algorithm by replicating previous work in sparse aperture design based on optimization.

- Use the layout optimization algorithm to design apertures that overcome or are more resistant to the issues present in sparse aperture remote sensing systems. These include:
  - The anisotropic nature of sparse aperture optical transfer functions.
  - The presence of ringing artifacts as a result of post-processing filters.
- Identify the effects of spectral knowledge on sparse aperture image quality after post-processing.
- Analyze the effects of narrow bandpass designs and broad bandpass designs on image quality under different spectral knowledge levels.
- Design and execute a laboratory experiment to validate expected spectral behavior in a simple imaging scenario.

The methods and results of the optimization study will be presented in Chapter 7. Due to the similarity of the analysis, the spectral bandpass results will be presented alongside the artifact validation results in Chapter 6.

Chapter 3

Background

This chapter will lay out the theory behind the prediction of the performance of imaging systems with non-conventional aperture functions. In order to fully understand the modeling of these systems, theory will be presented on the majority of the imaging chain. The main differences for non-conventional aperture systems are confined to light acquisition and image processing. However, a full knowledge of the imaging chain is necessary to construct an accurate image quality model. A reader interested in further imaging chain analysis in the context of remote sensing is referred to [Schott, 2007].

The theoretical modeling approach taken will rely heavily on the linear systems theory for image formation [Gaskill, 1978]. This approach allows for the use of convolution to model the degradation of imagery due to diffraction, aberrations and image motion. This framework also allows for the introduction of noise into the system and the derivation of the linear filters to apply in post-processing image enhancement. Aberration theory is also necessary to understand the various effects other than diffraction that can degrade image quality. The radiometry associated with such an imaging system will also be introduced.

3.1 Linear Systems in Imaging

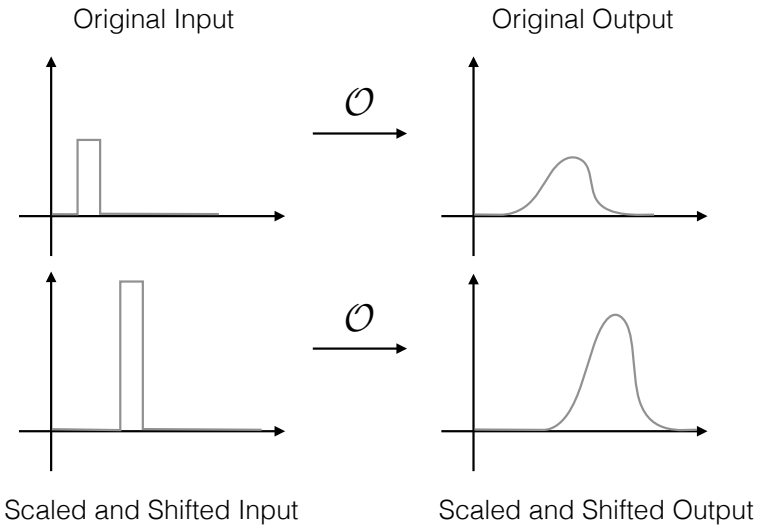


Figure 3.1: Illustration of linearity and shift invariance

Under the common assumption that an imaging system is both linear and shift invariant, linear systems theory may be used to predict the performance of that system. If we assume that the imaging process is some operator, $\mathcal{O}$, then the process is linear if

$$\mathcal{O}\{a \cdot f_1[x,y] + b \cdot f_2[x,y]\} = a \cdot \mathcal{O}\{f_1[x,y]\} + b \cdot \mathcal{O}\{f_2[x,y]\} \tag{3.1}$$

where a and b are scalars and $f_n[x,y]$ is an input to the imaging system. The imaging operator $\mathcal{O}$ is said to be shift invariant if

$$g[x,y] = \mathcal{O}\{f[x,y]\} \implies g[x-x_0, y-y_0] = \mathcal{O}\{f[x-x_0, y-y_0]\}$$

where x_0 and y_0 are scalar constants. This property is illustrated in Figure 3.1. As shall be seen later, the assumption of linearity and shift invariance is normally invalid across the entire image plane, however, it can be made over small portions of the image plane and thus linear systems theory can be used in a piecewise manner over the image plane to predict the performance of an imaging system.

Given that a system is both linear and shift invariant (LSI), the imaging operator, $\mathcal{O}$,

may be characterized as a convolution. If the input to the system is a point source of light, represented as a Dirac delta function, then the output predicted by the convolution will be the convolution kernel for the imaging system, $h[x, y]$. Since this kernel describes the image of a point source, it is termed the Point Spread Function (PSF) of the imaging system. The performance of the imaging system is then characterized in an LSI region by

$$\begin{aligned} g[x,y] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h[\alpha,\beta]\cdot f[x-\alpha, y-\beta]\ d\alpha\ d\beta \qquad (3.2)\\ &= f[x,y] * h[x,y] \qquad (3.3)\end{aligned}$$

When dealing with imaging system modeling, these convolutions are often computed in the frequency domain through the use of the Fourier transform operator, $\mathcal{F}$, defined as

$$F[\xi,\eta] = \mathcal{F}\{f[x,y]\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f[x,y]e^{-2\pi i(\xi x+\eta y)}\ dx\ dy \qquad (3.4)$$

The advantage of utilizing the frequency domain is that convolutions, which are computationally expensive to evaluate in the spatial domain, can be evaluated as simple element-wise multiplications in the frequency domain, due to the Fourier convolution theorem [Easton, 2010]. Thus,

$$g[x,y] = f[x,y] * h[x,y] = \mathcal{F}^{-1}\{\mathcal{F}\{f[x,y]\}\cdot\mathcal{F}\{h[x,y]\}\} \qquad (3.5)$$

where $\mathcal{F}^{-1}$ is the inverse Fourier transform, defined as

$$f[x,y] = \mathcal{F}^{-1}\{F[\xi,\eta]\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F[\xi,\eta]e^{+2\pi i(\xi x+\eta y)}\ d\xi\ d\eta \qquad (3.6)$$

This property is useful due to the existence of the Fast Fourier Transform (FFT), an algorithm to evaluate the Fourier transform in $O(n\log n)$, or log-linear, time. Since evaluating the two forward and one inverse transform all take log-linear time and the element-wise multiplication is a $O(n)$, or linear time, operation, the entire evaluation is a log-linear, $O(n\log n)$, computation, where n is the number of pixels in the image. On the other hand, if two arrays of size n need to be convolved in the spatial domain, one array needs to be

shifted to be centered at every element of the other array and the two then need to be multiplied and summed at every shift. This results in a quadratic, $O(n^2)$, time process. Thus, for reasonably large arrays, the frequency domain evaluation will be more efficient than the spatial domain evaluation. An example of this tradeoff is given in Figure 3.2.

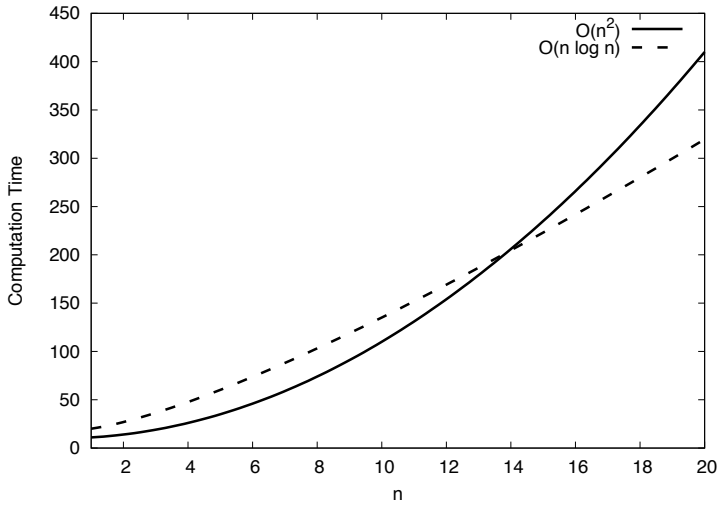


Figure 3.2: Example of the computation time curves for $O(n \log n)$ vs $O(n^2)$ algorithms. Precise values of n and computation time depend on the implementation.

Given the Fourier transform and the convolution theorem, the imaging relationship given in Equation 3.3 can be re-expressed as

$$G[\xi, \eta] = F[\xi, \eta] \cdot H[\xi, \eta] \tag{3.7}$$

where $H[\xi, \eta]$ is the Fourier transform of the system's PSF. This term is commonly referred to as the Optical Transfer Function (OTF) of the imaging system. This formulation of the imaging problem is highly advantageous for imaging system modeling when using the Imaging Chain approach.

While Equation 3.7 is very simple, it does not take into account noise in the system. Noise is inherent in every imaging situation. Shot or photon noise is inherent due to the Poisson nature of arriving photons and noise is also added by the detection mechanisms

of the system. Linear systems theory often makes the simplifying assumption that noise is additive and independent of the signal, so that Equation 3.7 may be rewritten as

$$G[\xi, \eta] = F[\xi, \eta] \cdot H[\xi, \eta] + N[\xi, \eta] \tag{3.8}$$

While this assumption is invalid due to photon noise, it is only utilized when deriving a linear image restoration filter, as will be discussed later. As a result, the final image quality is sub-optimal due to this assumption and a better, non-linear, restoration procedure may exist.

3.2 The Imaging Chain Approach

As seen in Figure 3.3, the Imaging Chain approach looks at the imaging process as a series of steps. The process begins at the light source and then considers interactions with the object of interest and its surroundings. Light then continues to interact with the environment until some of it reaches the imaging system's light collection apparatus. Once light is collected, it must be detected by the imaging system and converted into a measurable signal. This signal is then processed to produce a final image. This final image may then be displayed and perceived by an end user or automatically analyzed by a computer.

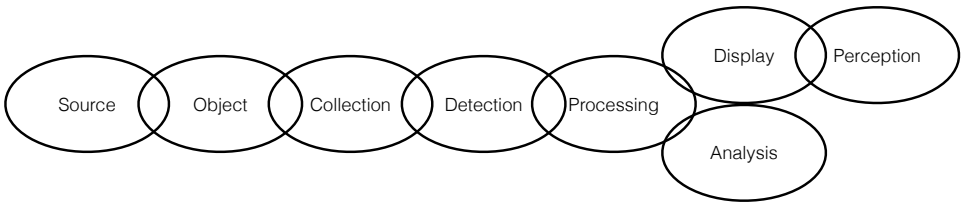


Figure 3.3: The Imaging Chain approach breaks down the imaging process into a series of links with limited interactions.

Each of these steps in the imaging chain is not perfect and has the potential to degrade or change the signal. For instance, light transport in the object link might blur the signal due to atmospheric scattering. It can also lower the signal level due to absorption. In this research, the focus will be on optical systems in the collection link. These can also blur the signal through diffraction or aberrations. Detection systems will also degrade the signal.

Modern digital detectors have a pixel footprint, or a rectangular region over which photons are integrated. This is a blurring and sampling function of the light distribution that was imaged onto the detector and thus a signal degradation. Once images are captured, image enhancement algorithms are performed on the image, again changing the signal, although ideally not a degradation. The Imaging Chain approach ties nicely into the linear systems approach described in the previous section if linear shift invariant assumptions are made on each of the steps. In this case,

$$G[\xi,\eta] = \left(\prod_i H_i[\xi,\eta]\right) \cdot F[\xi,\eta] \tag{3.9}$$

where $H_i[\xi,\eta]$ is the OTF produced by each step in the imaging process. Note that each link in the imaging chain is complex and can contain multiple steps.

Since light-object interactions are invariant with respect to the design of passive imaging systems, this research will be focusing on the light collection, detection and processing steps. Every optical component in an imaging system has the ability to degrade the signal and thus each optical element has its own OTF. The detector also has an associated OTF. In addition to the sampling footprint, pixel crosstalk, or signal leakage, can further complicate the OTF of an imaging detector. Finally, signal post-processing, when done with linear filters, also have associated OTFs. The cascading of multiple OTFs, described in Equation 3.9, will be used extensively in the modeling of complex optical systems.

3.3 Propagation of Light to the Entrance Pupil

Before describing how an optical system will degrade image quality due to both diffraction and optical aberrations, it is necessary to examine how the signal gets to the optical system. This problem has been explored in depth in the remote sensing literature and the approach taken by [Schott, 2007] will be used in this research. Optical transfer functions are applied to the signal between the system's pupil planes and as such, the problem of how light propagates to the system's entrance pupil is of interest here. It has been shown

that for a Lambertian object, this radiance distribution can be approximated as

$$\begin{aligned} L(\lambda) &= E_{\mathrm{s,exo}}(\lambda)\cos\sigma'\tau_1(\lambda)\tau_2(\lambda)\frac{\rho(\lambda)}{\pi} + \epsilon(\lambda)L_{\mathrm{emis}}(\lambda,T)\tau_2(\lambda) + FE_{\mathrm{ds}}(\lambda)\tau_2(\lambda)\frac{\rho(\lambda)}{\pi} \\ &\quad + FE_{\mathrm{de}}(\lambda)\tau_2(\lambda)\frac{\rho(\lambda)}{\pi} + (1-F)[L_{\mathrm{bs}}(\lambda) + L_{\mathrm{be}}(\lambda)]\tau_2(\lambda)\rho(\lambda) \\ &\quad + L_{\mathrm{us}}(\lambda) + L_{\mathrm{ue}}(\lambda) \end{aligned} \tag{3.10}$$

In this equation, $E_{\mathrm{s,exo}}$ represents the exoatmospheric solar irradiance, σ' is the solar angle with respect to the target, τ_1 is the transmission from the top of the atmosphere to the target, τ_2 is the transmission from the target to the imaging system, ρ is the diffuse target reflectance, ϵ is the target's emissivity, L_{emis} is the blackbody radiance due to the target's temperature, T, F is the sky fraction of the hemisphere above the target, E_{ds} is the reflected downwelled irradiance, E_{de} is the emitted downwelled irradiance, L_{bs} is the average radiance reflected by non-sky background, L_{be} is the average emitted radiance from non-sky background, L_{us} is reflected upwelled radiance and L_{ue} is the emitted upwelled radiance.

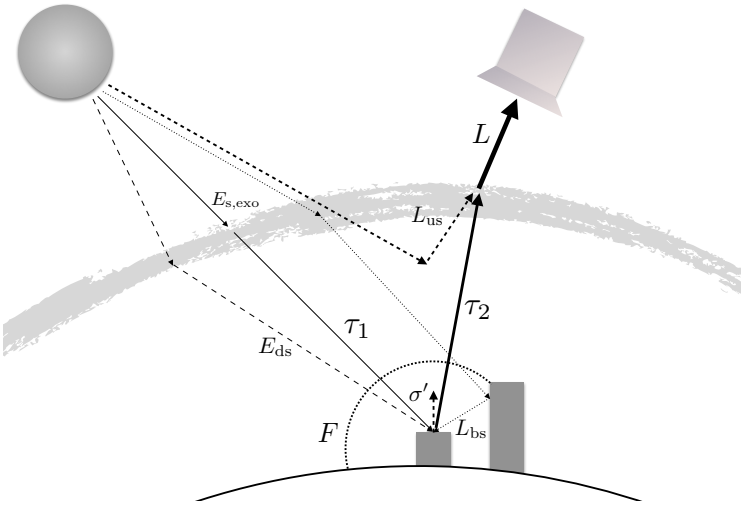


Figure 3.4: Illustration of the components of reflected sensor-reaching radiance.

Equation 3.10 gives the spectral radiance distribution reaching the imaging system due to reflection and emission. However, if the imaging system only detects in the visible to

shortwave infrared spectral regions, thermal emission is negligible, yielding,

$$\begin{aligned} L(\lambda) &= \left(E_{\mathrm{s,exo}}(\lambda)\cos\sigma'\tau_1(\lambda) + FE_{\mathrm{ds}}(\lambda)\right)\frac{\rho(\lambda)\tau_2(\lambda)}{\pi} \\ &\quad +(1-F)L_{\mathrm{bs}}\tau_2(\lambda)\rho(\lambda) + L_{\mathrm{us}} \end{aligned} \tag{3.11}$$

The components of this equation are visualized in Figure 3.4. In these equations, the transmission terms, downwelled irradiance and upwelled radiance are all atmospheric terms that need to be modeled. Traditionally, this is done using an atmospheric models, such as MODTRAN [Berk et al., 1989]. The sky fraction and average background radiance terms depend on the scene structure around the target. The radiance field reaching the front of the system predicted by Equation 3.11, represents the ideal image that can be obtained through the imaging system. Given that no imaging system can perfectly replicate this radiance field, the degration effects of the optical system on image quality must now be examined.

3.4 Diffraction-Limited Imaging

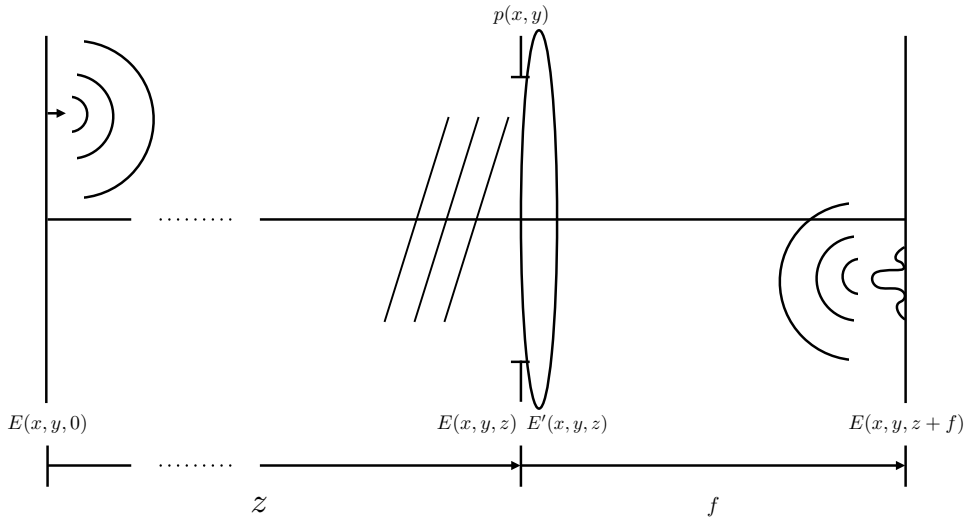


Figure 3.5: Imaging through a perfect lens with a finite aperture

As was alluded to in the Imaging Chain section, each optical element in a system will degrade the quality of the image. Every optical component, even one made perfectly to specification, will degrade the signal due to the fact that it has a finite size. The finite size of the optic results in diffraction effects, which blur the light distribution. An optic which degrades the signal primarily due to its diffraction effects and not any imperfections in the optic itself is termed diffraction-limited. To see the effects of this, an example of a finite lens with perfect focusing ability will be examined. This is shown in Figure 3.5. In this figure, the light distribution at the object plane, $z = 0$, is shown as an off-axis point source, the light then propagates over a long distance z to the lens. As derived in [Easton, 2010], Fresnel diffraction is used to propagate the light to the lens and from the lens to the image plane. The aperture function attenuates the wavefront at the aperture plane. The expression for the final electric field at the image plane is actually a convolution with the input signal, where the impulse response is given by,

$$\hat{h}\left[x, y; z_1 = \left(\frac{1}{f} - \frac{1}{z_0}\right)^{-1}\right] \propto P\left[\frac{x}{\lambda z_1}, \frac{y}{\lambda z_1}\right] \tag{3.12}$$

where P is the Fourier transform of the pupil function. This is an important observation; in the absence of all other aberrations, the degradation in image quality by an optic is determined by the shape of the aperture function. However, this expression operates on the electric field of the light, which is not directly measurable by modern imaging detectors. Instead these detectors measure exposure, i.e. irradiance integrated over a finite time interval. Irradiance is defined as the time average of the squared magnitude of the electric field. So, in this case, the PSF due to diffraction, normalized to unit area, is given by

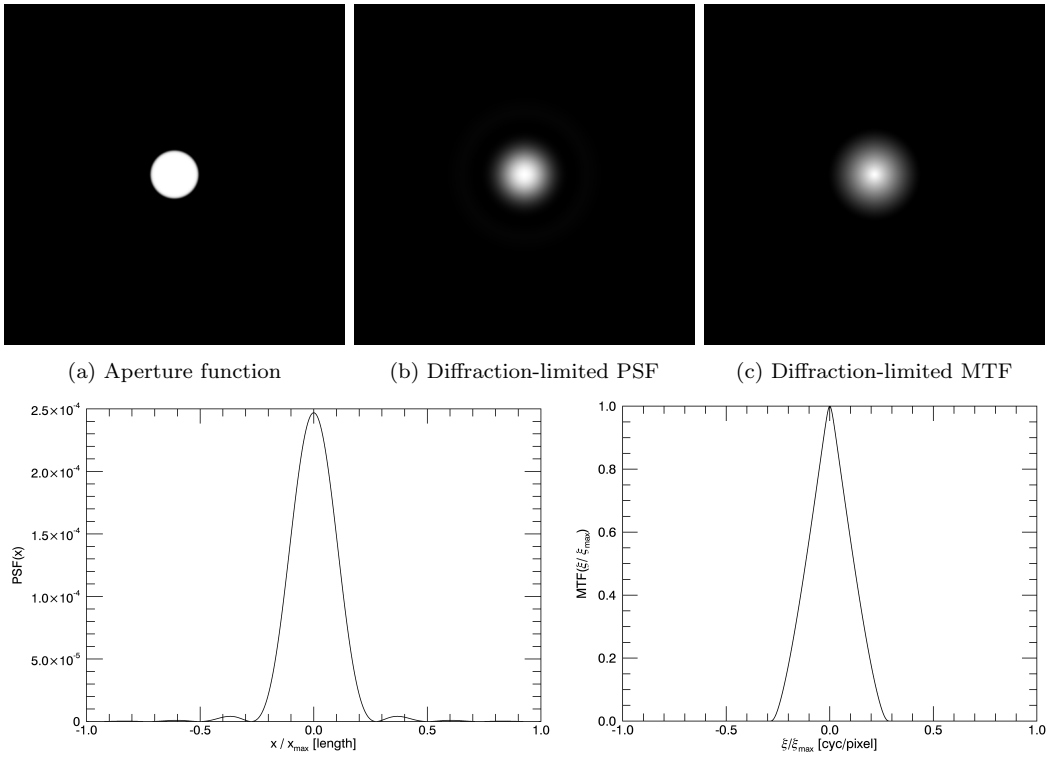
$$h\left[x, y; z_1 = \left(\frac{1}{f} - \frac{1}{z_0}\right)^{-1}\right] = \frac{1}{k}\left|P\left[\frac{x}{\lambda z_1}, \frac{y}{\lambda z_1}\right]\right|^2 \tag{3.13}$$

$$k = \iint_{-\infty}^{\infty} \left|P\left[\frac{x}{\lambda z_1}, \frac{y}{\lambda z_1}\right]\right|^2 dx\, dy \tag{3.14}$$

The OTF due to diffraction is then given by

$$H\left[\xi, \eta; z_1 = \left(\frac{1}{f} - \frac{1}{z_0}\right)^{-1}\right] = \frac{1}{k} \cdot p[-\lambda z_1 \xi, -\lambda z_1 \eta] \bigstar p[-\lambda z_1 \xi, -\lambda z_1 \eta] \tag{3.15}$$

where the ★ operator signifies the autocorrelation operation and p is the spatial domain pupil function.



(a) Aperture function (b) Diffraction-limited PSF (c) Diffraction-limited MTF

(d) 1D-profile of the PSF (e) 1D-profile of the MTF

Figure 3.6: Diffraction-limited imaging performance of a circular aperture.

3.4.1 Symmetric Aperture Functions

The performance predicted by the diffraction-limited assumption applies equally to systems with symmetric and non-symmetric aperture functions. As many non-symmetric designs utilize multiple symmetric designs as "building blocks", it is useful to examine the behavior of symmetric apertures. The simplest symmetric aperture is the unobstructed circular aperture. According to Equation 3.14, the associated diffraction-limited PSF is given by the squared magnitude of the Fourier transform of the circle function. As shown in Figure 3.6, this is given by the "Airy Disk" pattern, also known as a squared Bessel Sinc function [Easton, 2010].

While the circular aperture is a simple case, it is not always practical to build space-based telescopes with a circular aperture. Many of these telescopes instead have a secondary focusing mirror that blocks the central portion of the primary mirror, as occurs in the Cassegrain design, shown in Figure 3.7. Since the secondary mirror blocks some rays from entering the system, it introduces a central obscuration into the aperture function. This obscuration has little effect on the low- or high-frequency values of the MTF, however, it does cause a dip in the mid-frequency response of the system. In the spatial domain, the peak of the PSF has been reduced in magnitude. More energy is also in the secondary peaks than in the case of the unobstructed circular aperture. These effects can be seen in Figure 3.8.

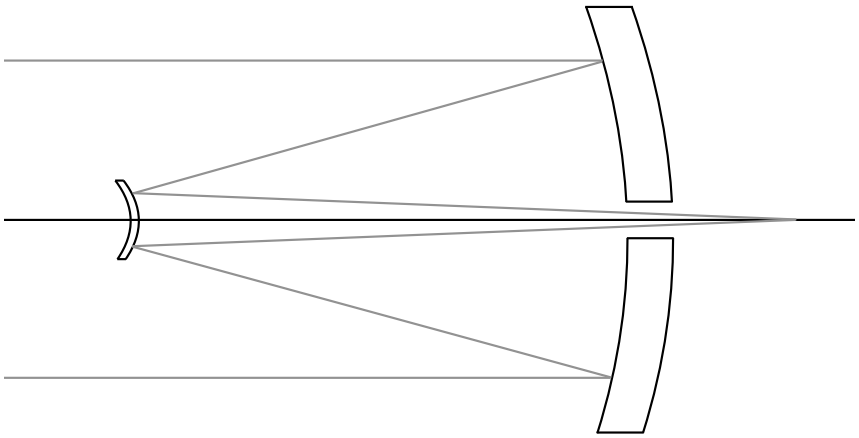
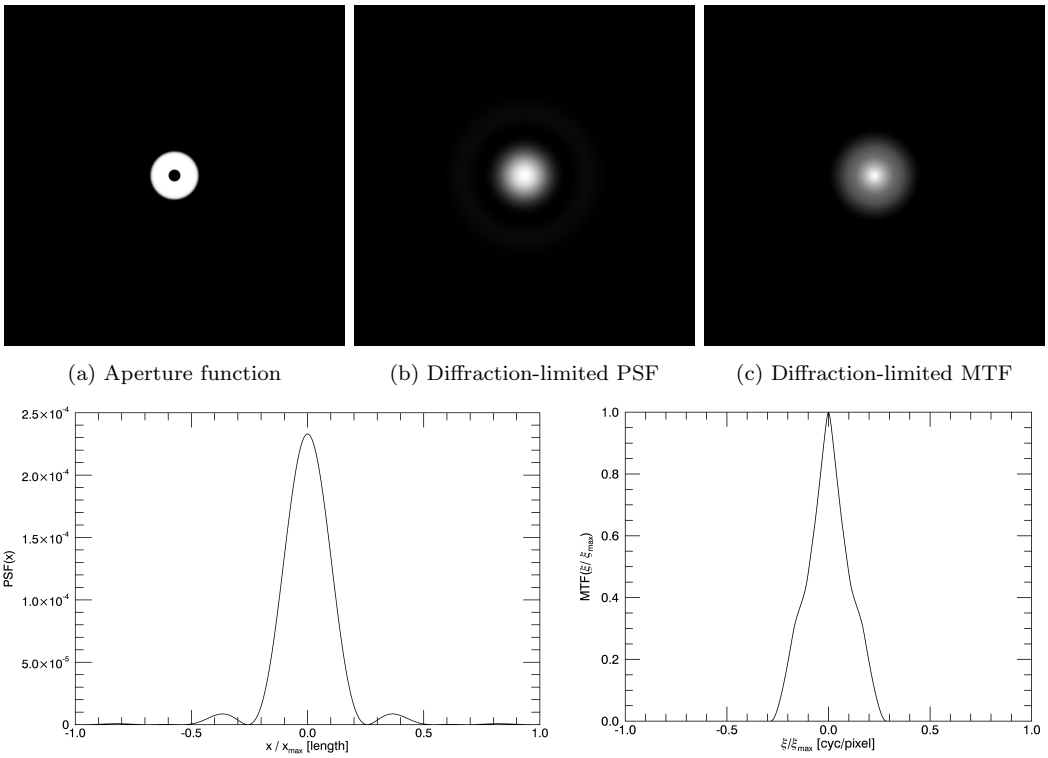


Figure 3.7: The Cassegrain telescope design



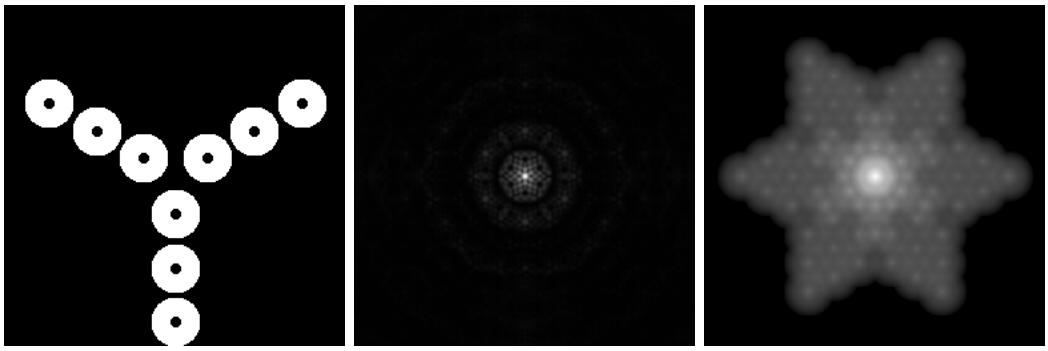
(a) Aperture function (b) Diffraction-limited PSF (c) Diffraction-limited MTF

(d) 1D-profile of the PSF (e) 1D-profile of the MTF

Figure 3.8: Diffraction-limited imaging performance of an obstructed circular aperture.

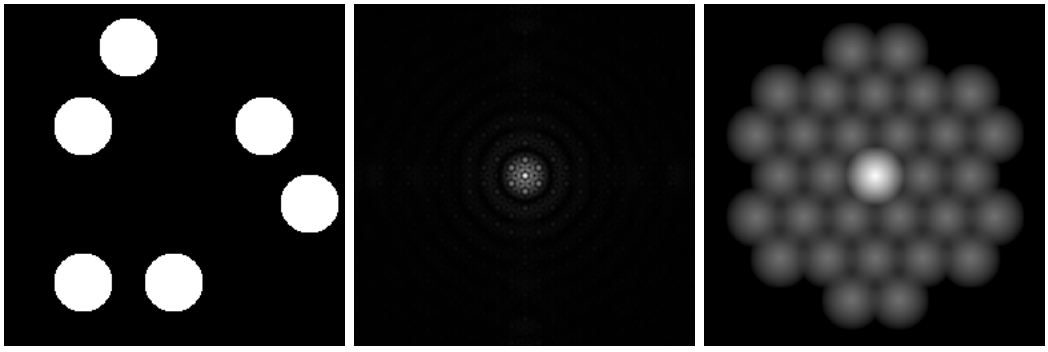
3.4.2 Non-symmetric Aperture Functions

The aperture functions discussed in the previous section were all symmetric about the origin and have been widely studied. However, the same analysis can be applied to systems that do not have this symmetry. These non-symmetrical systems have the potential to have useful imaging capabilities, such as obtaining more resolution using the same or lesser amounts of glass. Figure 3.9 shows the MTF and PSF associated with a Tri-arm 9 sparse aperture design. The Tri-arm design is commonly seen in sparse aperture studies, such as [Fiete et al., 2002]. This is due to its proven effectiveness through its use in the Very Large Array radio telescope (VLA) in Socorro, NM [Napier et al., 1989]. Figure 3.10 shows the MTF and PSF associated with the Golay-6 design. This design was theorized in [Golay, 1971] and gives optimal frequency coverage and MTF compactness. The Golay-6 design will be heavily in the laboratory experiments in this research.



(a) Aperture function (b) Diffraction-limited PSF (c) Diffraction-limited MTF (contrast-stretched)

Figure 3.9: Diffraction-limited imaging performance of a Tri-arm 9 design.



(a) Aperture function (b) Diffraction-limited PSF (c) Diffraction-limited MTF (contrast-stretched)

Figure 3.10: Diffraction-limited imaging performance of a Golay-6 design.

This sparse aperture design exhibits some characteristics not shown in the symmetric apertures. Since the aperture function is non-symmetric, the PSF and MTF are also non-symmetric. This means that the resolution limit will vary based on the object's orientation. The MTF of the symmetric apertures seen in the previous section were monotonically decreasing with increasing spatial frequency. This is not the case, however, with this sparse aperture design, as the MTF shows multiple peaks. Each peak arises when two of the Cassegrain-style sub-apertures overlap in the calculation of the autocorrelation function. Since the Tri-arm 9 design is an array of identical sub-apertures, the sifting

property of the Dirac delta function can be used to reformulate the pupil function,

$$p[x,y] = \sum_{i=1}^{N} p_{\text{sub}}[x-x_i, y-y_i] \tag{3.16}$$

$$= p_{\text{sub}}[x,y] * \sum_{i=1}^{N} \delta[x-x_i, y-y_i] \tag{3.17}$$

where p_{sub} is the aperture function of the sub-aperture, (x_i, y_i) is the center position of the ith sub-aperture and N is 9 in this case. The diffraction limited OTF is then given by equation 3.15.

$$\begin{aligned} OTF &= \frac{1}{k} p_{\text{sub}}[-\lambda z_1 \xi, -\lambda z_1 \eta] * \sum_{i=1}^{N} \delta[-\lambda z_1 \xi - x_i, -\lambda z_1 \eta - y_i] * \\ &\quad p^*_{\text{sub}}[-\lambda z_1 \xi, -\lambda z_1 \eta] * \sum_{i=1}^{N} \delta[-\lambda z_1 \xi + x_i, -\lambda z_1 \eta + y_i] \\ &= \frac{1}{k} (p_{\text{sub}}[-\lambda z_1 \xi, -\lambda z_1 \eta] \bigstar p_{\text{sub}}[-\lambda z_1 \xi, -\lambda z_1 \eta]) * \\ &\quad \sum_{i=1}^{N} \sum_{j=1}^{N} \delta[-\lambda z_1 \xi - x_i + x_j, -\lambda z_1 \eta - y_i + y_j] \end{aligned} \tag{3.18}$$

This result shows that the diffraction-limited OTF of a sparse aperture design, composed of identical sub-apertures, is given by the sum of shifted copies of the diffraction-limited OTFs of the sub-apertures, which is why sparse aperture OTFs often exhibit secondary peaks.

3.5 Aberration Theory

Simply considering the diffraction due to the shape and size of the optics is not sufficient to accurately predict their performance. The assumption of diffraction-limited performance can be made only for optics that have been precisely constructed and aligned. Image degradations that occur from imperfections in the construction or alignment of the optical systems not due to diffraction are termed aberrations. In an ideal imaging scenario, wavefronts converge spherically to infinitesimal points. Deviations from this ideal wavefront are the cause of these aberrations, as shown in Figure 3.11.

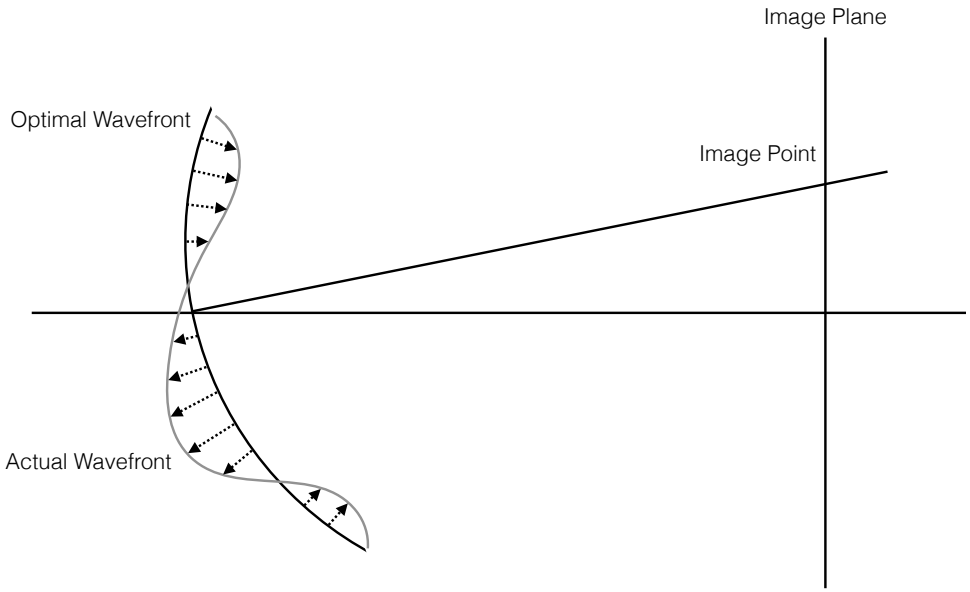


Figure 3.11: Deviations from a spherical wavefront result in aberrations

In the case of perfect imagery, the wavefront error is zero for all points along the wavefront. However, wavefront error tends to be non-zero and gets larger as the distance away from the optical axis increases. Just as the radiance field entering an optical system is a four-dimensional function, the wavefront aberration function is also four-dimensional. That is, it depends on both the two-dimensional coordinates on the aperture and the incoming direction. Equivalently, if there is a defined object plane, wavefront aberration depends on both aperture coordinates and object coordinates. This is illustrated with a sample aperture in Figure 3.12.

In the case of a circularly symmetric system, the dimensionality can be reduced by one. Since the aperture is circularly symmetric, the aberration is constant with respect to object orientation, so only the case of an axis-aligned object needs to be considered. Then, due to the symmetry, the aberration is dependent on the radius in aperture space, ρ and the cosine of the angle from the object's axis to the ray intersection on the aperture plane, $\cos\phi$. Under these conditions, aberration can only take on a limited number of forms, which are given up to fourth-order terms in Table 3.1. Some example wavefront error functions are shown in Figure 3.13.

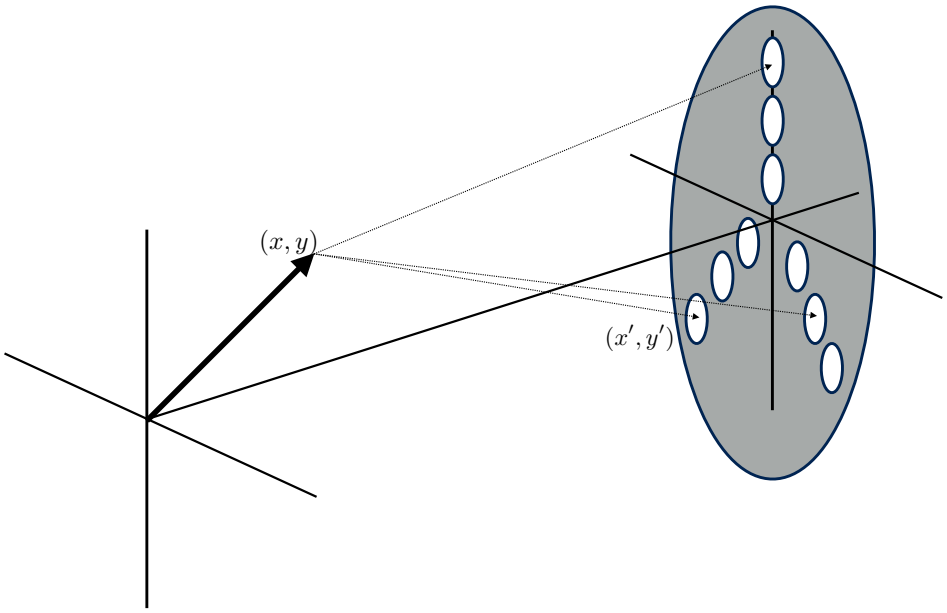


Figure 3.12: Wavefront aberration is a four-dimensional function, depending on the incoming ray direction and intersection coordinates on the aperture

The wavefront aberration function represents the change in optical path length based on the entrance location and angle of the optical path through the optical system. So, the aperture function can be re-expressed at a given object plane location, as

$$p[x, y, x', y'] = |p[x, y, x', y']|e^{2\pi i W(x,y,x',y')} \tag{3.19}$$

where $|p[x, y, x', y']|$ is the mask of the aperture, generally a binary "zero-and-one" function [Goodman, 2005]. It can be used to give the transmission through the optics for every optical path, however, a constant transmission is normally assumed and applied later. Note that the relationship between this complex aperture function, also known as the pupil function, and the OTF of the system still holds, as per Equation 3.15. So, the system's OTF now depends on the location in the object plane. Due to this dependence, the degraded image can no longer be computed as a simple convolution. However, the problem is still tractable if it is assumed that the wavefront error can be approximated by a piecewise constant with respect to location on the object plane or its conjugate location on the image plane. This gives rise to regions on the image plane that share a system

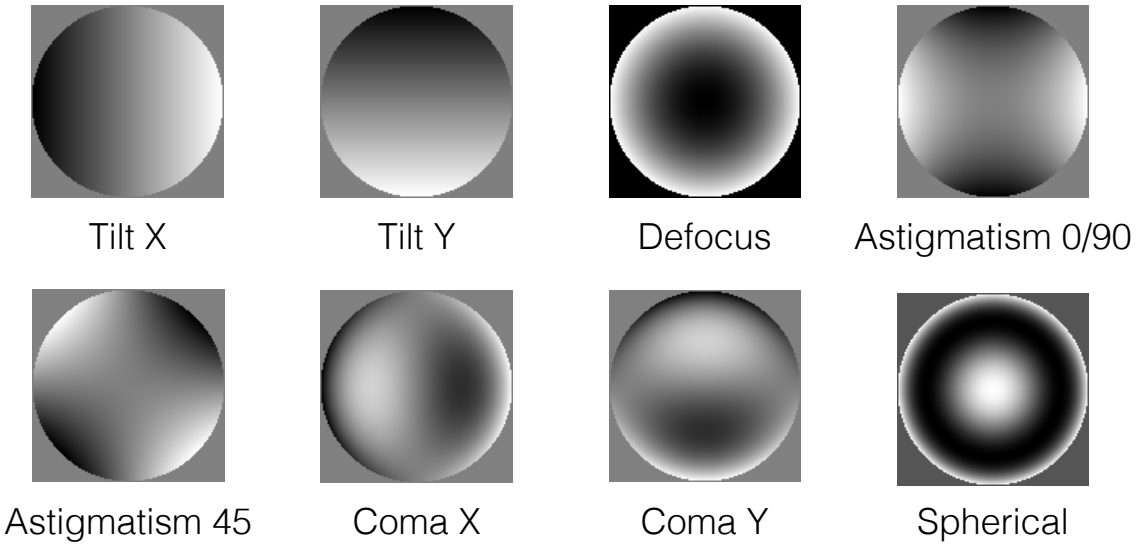


Figure 3.13: Example instantiations of wavefront aberrations. Brightness is proportional to optical path length error.

OTF, termed isoplanatic regions.

Table 3.1: Aberrations for circularly symmetric optical systems

Aberration	Wavefront Error $W(h, \rho, \cos\phi)$
Piston Error	$W_{000},\ W_{200}h^2,\ W_{400}h^4$
Defocus	$W_{020}\rho^2$
Tilt	$W_{111}h\rho\cos\phi$
Spherical	$W_{040}\rho^4$
Coma	$W_{131}h\rho^3\cos\phi$
Astigmatism	$W_{222}h^2\cos^2\phi$
Field Curvature	$W_{220}h^2\rho^2$
Distortion	$W_{311}h^3\rho\cos\phi$

3.6 Propagation of Light to the Sensor

After accounting for signal degradations by the optics due to aberration and diffraction, the radiance field given by Equation 3.10 then needs to be propagated to the detector.

As the previous sections have stated, there is a blurring effect that occurs between the entrance pupil and the exit pupil. In an isoplanatic region, this blurring can be evaluated as a convolution. In order to convert radiance reaching the exit pupil to irradiance on the detector, the G-Number ($G\#$) is used.

$$G\#(\lambda) = \frac{1 + 4(f/\#)^2}{\pi F \tau(\lambda)} \tag{3.20}$$

where $f/\#$ is the system F-Number, F is the system's fill factor, or the fraction of the pupil area that is transmissive, and τ is the transmission spectrum of the optics. The fill factor term is normally omitted in systems with circular apertures as it is equal to one, however, it must be included in systems with sparse apertures. The spectral irradiance on the detector in an isoplanatic region is then given as,

$$E_{\text{det}}(x, y, \lambda) = \frac{L_{\text{source}}(x, y, \lambda) * h(x, y, \lambda)}{G\#} \tag{3.21}$$

where $h(x, y, \lambda)$ is the point spread function given as the inverse Fourier transform of the OTF given in Equation 3.15. It should be noted that while it is not explicitly noted in the equations, L_{source} and thus E_{det} are random variables, due to photon noise. That is, they are actually Poisson variables, where the predicted radiometric value is both the mean and variance.

If it is assumed that the detector acts linearly at the input signal level, then the signal in volts for a given pixel on the detector is given as

$$S_{\text{volt}}(x, y) = A_d t_{\text{int}} \int_0^\infty E_{\text{det}}(x, y, \lambda) R(\lambda)\ d\lambda + N(x, y) \tag{3.22}$$

where A_d is the area of the pixel, t_{int} is the integration time of the detector and $R(\lambda)$ is the responsivity spectrum of the detector and $N(x, y)$ is the noise added by the detector. $N(x, y)$ consists of all detector-generated noise sources, including dark current and read noise. Modern digital detectors have an analog-to-digital converters which convert a voltage from the signal in electrons to a digital count that gets recorded to form the final image. This will be examined in the detector modeling section in the next chapter.

3.7 Image Enhancement

As was seen in Figure 3.9, a sparse aperture function can result in an OTF that has support, i.e. energy, over a wide range of frequencies, even if the magnitude at those frequencies is relatively low. If the noise in the system is also relatively low, then the images produced by these systems are good candidates for image enhancement. Inverse filtering is the operation of trying to undo the process of a convolution, retrieving the input signal. Using an additive noise model and a convolution to model image degradation, the output image spectrum is given by

$$G[\xi,\eta] = F[\xi,\eta] \cdot H[\xi,\eta] + N[\xi,\eta] \tag{3.23}$$

The goal of inverse filtering is to recover F as closely as possible. Since noise is unknown and convolution is almost always a non-invertible operator, the problem must be approached as a minimization of some error between the original object and the reconstructed image. A common error metric is the sum of the squared error. That is, the reconstructed image, $\hat{f}$, satisfies

$$\underset{\hat{f}}{\operatorname{argmin}} \sum_{y=1}^{N}\sum_{x=1}^{M}(f[x,y] - \hat{f}[x,y])^2 \tag{3.24}$$

This problem is complicated by the fact that in an imaging situation, neither the object nor its spectrum is known. However, it can be shown that, given the assumption that the noise and signal are statistically independent, the minimum squared error can be minimized by the following filter function, known as the Wiener filter. [Easton, 2010]

$$W[\xi,\eta] = \frac{1}{H[\xi,\eta]} \frac{|F[\xi,\eta]|^2}{|F[\xi,\eta]|^2 + |N[\xi,\eta]|^2} \tag{3.25}$$

This equation can then be re-expressed in its more common form

$$W[\xi,\eta] = \frac{H^*[\xi,\eta]}{|H[\xi,\eta]|^2 + \frac{|N[\xi,\eta]|^2}{|F[\xi,\eta]|^2}} \tag{3.26}$$

An example of the Wiener filter is shown in Figure 3.14. This filter looks similar to that of a naive inverse filter, with the addition of the noise to signal power spectrum ratio in

the denominator. The effect of this term is to make the filter act as an inverse filter at frequencies where the signal-to-noise ratio is high and as a noise suppressor at frequencies where the signal-to-noise ratio is low. This overcomes the primary downfall of the naive inverse filter, which is the boosting of noise at spatial frequencies with low signal-to-noise ratios.

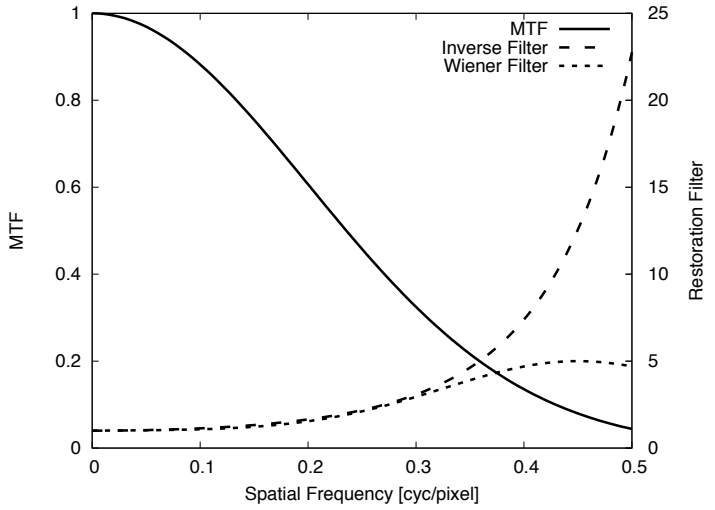
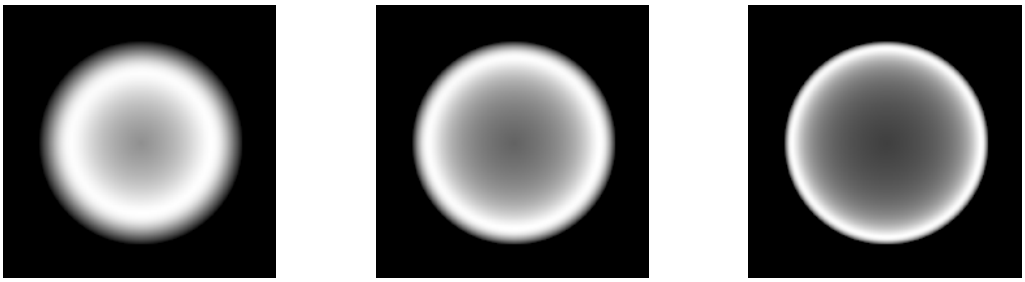


Figure 3.14: An example of a Wiener filter with a constant noise spectrum

As can be seen in Equation 3.26, the transfer function in the numerator is a complex function. This is advantageous as the transfer function of optical systems is often complex due to wavefront aberrations. If these wavefront aberrations have been characterized for a system, the system's complex transfer function can be input into the Wiener filter in order to compensate for these aberrations. This can be very useful in the case of sparse aperture systems. Small aberrations can result in the disappearance of some of the peaks in the system's OTF. Using the unaberrated OTF in the Wiener filter would result in inverse filtering at these frequencies. Since the aberrations eliminated the signal at those frequencies, the unaberrated Wiener filter would simply boost noise and introduce ringing artifacts, degrading the resulting image.

The Wiener filter, while minimizing the squared error, can often be impractical to use. The largest issue is that the power spectrum of the object is almost always unknown. A number of approaches can be taken here. If the noise power spectrum is known or ap-



(a) $\gamma = 0.1$ (b) $\gamma = 0.01$ (c) $\gamma = 0.001$

Figure 3.15: Examples of constrained least-squares restoration filters for a circular aperture. Notice that as γ decreases, higher frequencies are increasingly boosted.

proximated, the output image spectrum can be used as an approximation of the object spectrum. The filter can then be iterated, updating the image spectrum after each iteration. This is known as the iterative Wiener-Helstrom filter [Schott, 2007]. Another option is to simply use a constant value for the noise-to-singal term. The option used in this research was a constrained least-squares variant of the Wiener filter. [Reddi, 1978]

$$\hat{F}[\xi,\eta] = W[\xi,\eta] \cdot G[\xi,\eta] = \frac{H^*[\xi,\eta]}{|H[\xi,\eta]|^2 + \gamma \cdot |S[\xi,\eta]|^2} \cdot G[\xi,\eta] \tag{3.27}$$

Instead of simply using a constant in the denominator, a local image smoothness term is added to the denominator. This term $S[\xi,\eta]$ is the transfer function of a Laplacian convolution kernel. This term is modulated by a tunable scalar, γ that can be used to adjust the amount of inverse filtering or blurring that occurs as a result of the filter. The effect of this parameter on the restoration filter is shown in Figure 3.15. Optimal determination of this parameter is difficult, but depends on the noise present in the output image and the use case of the output image. For instance, the value of γ that is optimal for human perception at a given noise level might differ from the optimal value of γ if the image were to be used for some automated processing. In practice, the value of γ is empirically tuned for the given imagery to give an optimal output. For low noise inputs, γ tends to be less than one. As γ increases beyond one, the inverse filter becomes more of blurring operation to suppress noise than a sharpening filter.

Chapter 4

Modeling Approach

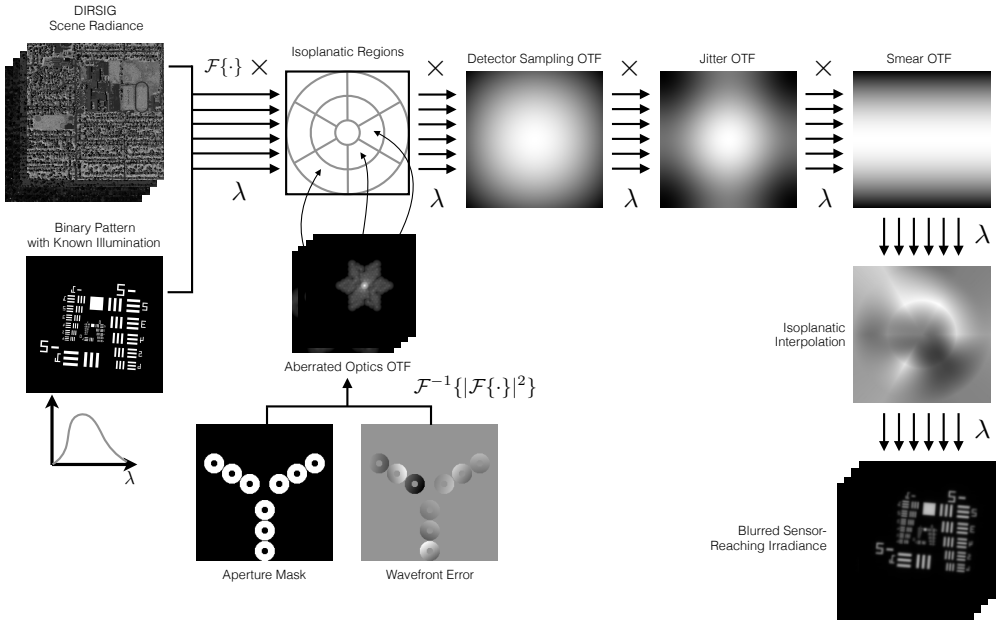


Figure 4.1: A flowchart visualization of the modeling approach to produce raw sensor-reaching irradiance.

This chapter will explain the approach taken to model optical systems with non-conventional aperture functions, applying the theory presented in the previous chapter. A graphical overview of the first half of the model is given in Figure 4.1. In this figure, the process

of obtaining the irradiance field incident on the detector is shown. The process by which this is converted into a final image is illustrated in Figure 4.2.

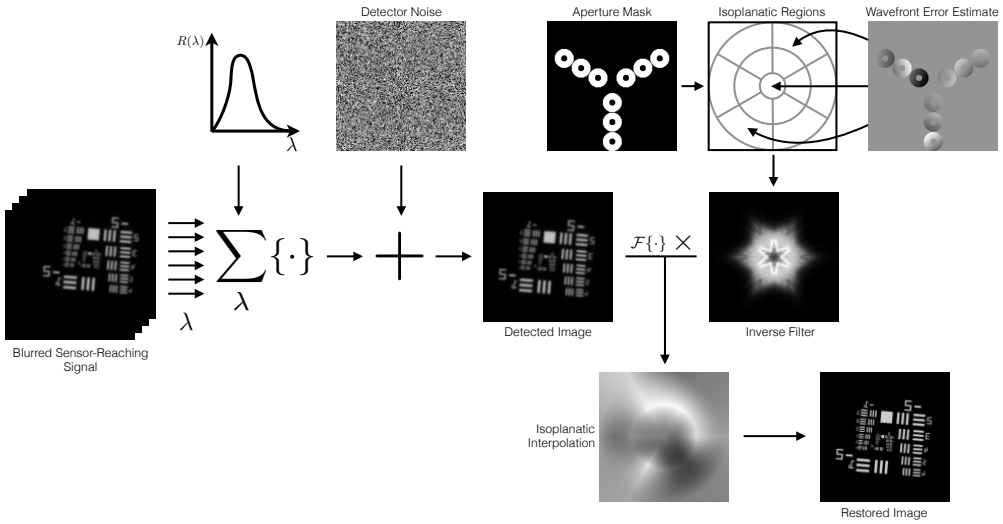


Figure 4.2: A flowchart visualization of the modeling approach to produce a final restored image.

Each element in these flowcharts will be examined in detail in the subsequent sections. In summary, there are two stages to the model that require input: imaging and restoration. When imaging, a spectral radiance image at the entrance pupil of the system is required. This can come from a number of sources, but in this research it is either generated by the Digital Imaging and Remote Sensing Image Generation (DIRSIG) model [Schott et al., 1992], or produced from a binary pattern with a constant illumination function. Photon noise is modeled and added to the input image in both cases. This spectral radiance image is then degraded using a series of optical transfer functions. The first is the spectrally-varying optical transfer function due to the aberrated optics of the system, which will require user input. The first input is a binary mask, which represents the shape of the sparse aperture array. The second input is the wavefront error at each point on the aperture. As was mentioned in Chapter 3, the wavefront error can vary with both position on the aperture and position on the image plane, which results in a four-dimensional functions. The model allows for this behavior through the use of isoplanatic regions and isoplanatic interpolation. After degradation by the optics, the image is then degraded with transfer functions due to the footprint of the detector, jitter in the optical system

and smear due to linear motion over the integration time, all of which are assumed to be constant over the spectral dimension.

Once degraded, the signal is spectrally summed over the bandpass(es) of the detector and degraded by the addition of detector noise, giving an approximation of Equation 3.22, or the raw image that was detected by the imaging system. This image is then processed by a restoration filter to produce the final restored image. This restoration filter is produced from knowledge of the aperture shape and an approximation of the wavefront error across the aperture, as it may be impractical to precisely measure this error once the system has been deployed. This inverse filtering can also be applied in a spatially-varying manner, if off-axis aberration in the system can be characterized.

4.1 Previous Work

Sparse aperture arrays have had a long history in both radio astronomy, infrared astronomy and optical physics. In 1970, a sparse aperture telescope for use in the long-wave infrared was being designed by [Meinel, 1970]. The problem of layout optimization also has a long history, with [Golay, 1971] designing aperture configurations to give desirable MTF characteristics. More recently, research has picked up into modeling and creating sparse aperture telescopes for remote sensing purposes. Unlike previous research into the topic, the area of interest here is in the visible to near-infrared region of the electromagnetic spectrum. This is a more challenging problem, as alignment tolerances scale with wavelength, resulting in the need for optical systems with an extreme amount of precision. [Fiete et al., 2002] from the Eastman Kodak Company have performed a number of image quality studies exploring the trade spaces of sparse aperture design. Confining analysis to three well-known sparse aperture designs (Tri-arm 9, Golay 6 and annulus), their work analyzed the tradeoff of fill factor and integration time, finding that integration time had to be increased by a factor in the range of $1/F^2$ to $1/F^3$, depending on aperture design. This conclusion agreed with the conclusion of [Fienup, 2000], who found the $1/F^3$ factor for the annulus design. Fiete et. al. also defined an "effective diameter" for a sparse aperture's MTF as the geometric mean of the minimum and maximum diameter of the MTF. Their work showed that NIIRS and GIQE were not applicable to sparse aperture designs.

[Breckinridge et al., 2008] built upon this work even further, looking at more optical designs and low-contrast imaging situations. In such situations, they found the exponent in the integration time versus fill factor tradeoff could reach 4 or 5 depending on contrast level and subaperture layout. The large effects of signal-to-noise ratio on the imaging parameters of sparse aperture systems in this work also strengthen the argument against using GIQE for sparse aperture imagery. In the GIQE, SNR is weighted relatively lightly compared to spatial resolution and the SNR levels of the conventional imagery used to fit the GIQE were obtained with a much shorter integration times than would be required in a sparse system.

The model outlined in the chapter introduction is the result of several research projects. The original model was introduced by [Introne et al., 2005]. In this work, the three well-known sparse aperture designs from [Fiete et al., 2002] were extensively analyzed using the modeling approach. This work also established the importance of using polychromatic simulation in sparse aperture modeling, showing that a grey-world assumption was not sufficient for image quality studies with these systems. This work was extended by [Block, 2005], who examined the spectral issues of sparse aperture imaging in more detail. This work conducted a sensitivity study to examine the nature of spectral artifacts that arise due to inverse filtering in a panchromatic system. The findings of these works are the primary focus of the validation study proposed in this research.

The modeling approach originally developed by Introne is not confined to sparse aperture systems. As [Zelinski and Schott, 2009] showed, the approach was also practical for modeling systems with synthetic apertures, such as the James Webb Space Telescope. This work also showed that the model was not only capable of handling rigid aberrations (piston, tip, tilt), but also aberrations due to warping of mirrors that can happen in foldable synthetic apertures. This work also noted that the NIIRS scale and GIQE were not ideal image quality metrics for exotic apertures, agreeing with the conclusion of Fiete et. al.. Instead, a more task-based approach was adopted for modeling quality. Specifically, a multispectral motion detection algorithm was used as a benchmark for image quality, where various changes were evaluated based on their effects on the results of the algorithm. Finally, this approach of task-based image quality was brought back to the realm of sparse

apertures by [Smith, 2012]. In this work, target detection was used as the benchmark algorithm with comparative analysis performed on monolithic, Tri-arm 9 and hexagonal synthetic apertures.

All of the above works were confined to computer modeling of sparse aperture performance. [Chung et al., 2002] from the Massachusetts Institute of Technology have built a ground-based three-element sparse aperture telescope, called ARGOS. They have published an analysis of how the PSF degrades with alignment error and a cost analysis. Unfortunately, the cost of the system is in the hundreds of thousands of dollars, making their design impractical for a laboratory model validation study as proposed here. Additionally, no images from the telescope have been published. [Zhou et al., 2009] from Beijing University of Technology have constructed a low-cost laboratory setup utilizing masks to simulate sparse apertures. This setup is the one from which the setup in this research draws the most inspiration. Their setup utilizes refractive lenses (and thus has a limited spectral range) and has limited control over wavefront error, two factors this work will attempt to improve upon. Additionally, their study did not perform quantitative comparison to modeled PSF/OTFs and was limited to the three well-known configurations. This works aims to expand upon both of these areas.

4.2 Radiance Image

As was described in Section 3.3, there are many terms that contribute to the radiance distribution that reaches the entrance pupil of an imaging system. In a remote sensing application, one has to consider direct solar illumination, diffuse illumination from the sky and energy scattered into the line of sight, amongst other terms. If infrared radiation is being examined, emitted radiation must also be considered. In order to model how a potential system will perform, realistic synthetic imagery needs to be generated using the theory presented.

The Digital Imaging and Remote Sensing Image Generation (DIRSIG) model has been in development for over 20 years to fill this task [Schott et al., 1992]. This model uses ray tracing to model light transport from first-principles. The DIRSIG model functions over a large range of wavelengths and modalities and will be used around the visible and near-

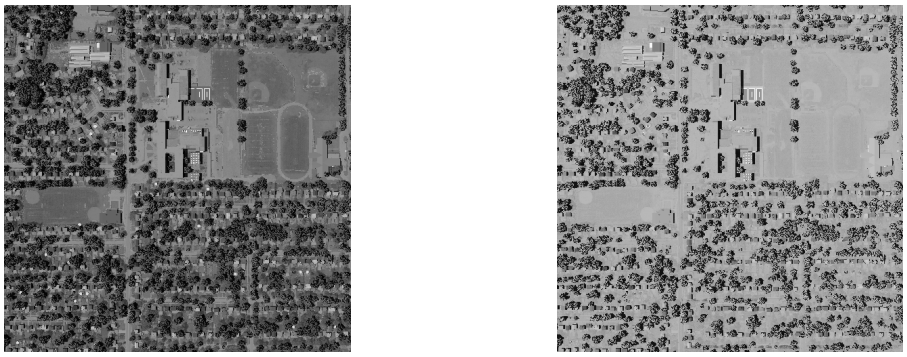
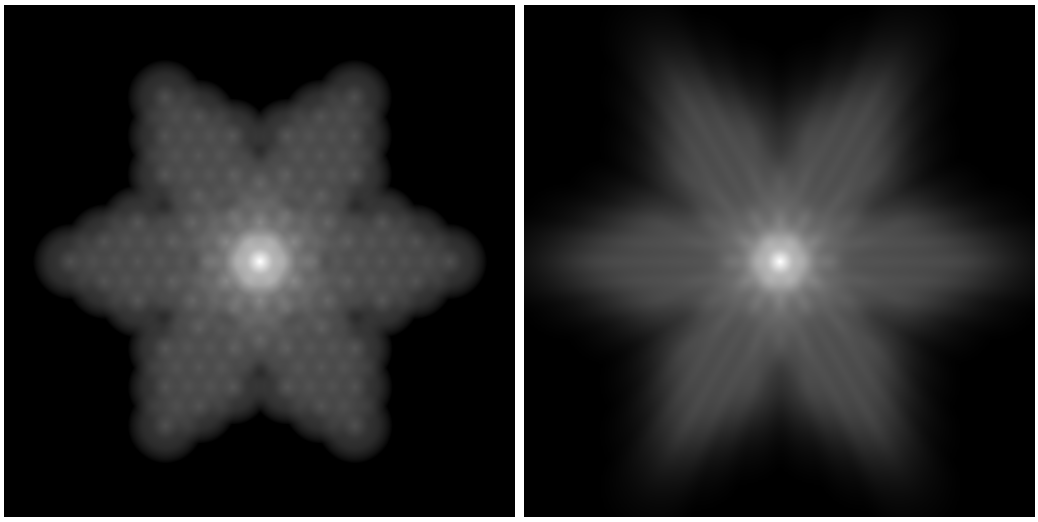


Figure 4.3: Sample bands of a DIRSIG-generated hyper spectral input to the model

infrared spectrum in this research. In order to produce high-fidelity images, the model needs to incorporate a lot of information about the scene. Some of these inputs include spectrally-attributed scene geometry, atmospheric conditions, scene thermal information and weather history. The technical details and intricacies of the model are out of the scope of this research, however, an interested reader is referred to the DIRSIG documentation (http://www.dirsig.org/docs/new/) for more details. In this research, the role of DIRSIG is to model the radiance distribution at the entrance pupil of the imaging system that we wish to model, that is, it is providing the $L_{\text{source}}(x, y, \lambda)$ term in Equation 3.21. Some example bands from a DIRSIG scene are given in Figure 4.3.

The radiance image, $L_{\text{source}}(x, y, \lambda)$, that is modeled by DIRSIG is a three-dimensional data source. It has two spatial dimensions and one spectral dimension, similar to a hyperspectral data cube. The model can be set up to vary the resolutions in all three of these dimensions. Since the DIRSIG model already performs spatial integration inside of a pixel through the use of adaptive sampling, the spatial resolution can be set to the size of the detector without any worry of introducing aliasing into the system, provided sufficient spatial oversampling was specified when generating the scene.

The resolution of the spectral dimension, however, can be more complicated to determine. Remote sensing systems have a very diverse set of output bands that might be modeled. For instance, panchromatic bands can have bandpasses that span the en-



(a) Center: 550 [nm], FWHM: 1.5 [nm] (b) Center: 550 [nm], FWHM: 150 [nm]

Figure 4.4: Illustration of the effects of wavelength on the diffraction-limited performance of a Tri-arm 9 sparse aperture design.

tire visible range, while hyperspectral systems can have bandpasses that span only a few nanometers. Thus, the spectral resolution of the input radiance image should be dependent on the bandpasses that are going to be modeled and should always be higher than the bands being modeled. As was shown in Equation 3.15, the optical transfer function due to diffraction is highly dependent on wavelength. As such, the degradation of the signal can vary significantly over a panchromatic or multispectral band. This effect is illustrated in Figure 4.4 for a Tri-arm 9 sparse aperture for a band centered around 550 nanometers with a full-width, half-max (FWHM) of 150 nanometers in comparison to the same band with a FWHM of 1.5 nanometers.

As can be seen in Figure 4.4, the spectral change in OTF over the bandpass is very significant. For this sparse aperture design, all of the peaks in the periphery of the OTF have been blurred over. As will be seen later, this presents challenges for applying inverse filtering to the raw image. However, in terms of configuring DIRSIG, Figure 4.4 should make it apparent that spectral oversampling is necessary for high-fidelity modeling of the performance of these systems. In practice, samples every 10 nanometers yield a high-quality estimate for a panchromatic bandpass.

4.3 Aberrated Optics OTF

Under normal operating conditions, the optics of a sparse aperture system provide the primary degradation of image quality. As such, the modeling approach for this degradation should be as accurate as possible. The inputs to the model that facilitate this modeling are the aperture mask that defines the unaberrated pupil function and the wavefront error at each point on the aperture. These inputs are given as two-dimensional images. This representation carries some inherent assumptions. As was illustrated in Figure 3.12, wavefront error is a four-dimensional quantity. The aperture mask is representative of the transmission through the optics along every optical path, which is also four-dimensional. Due to the large object distances present in remote sensing situations, the field of views of these systems tend to be very small. Thus, within an isoplanatic region of the image plane, the transmission and wavefront error are assumed to be constant with respect to incoming ray angle. This assumption is closely tied to that of local linearity and shift-invariance, allowing the computation of image degradation as a convolution.

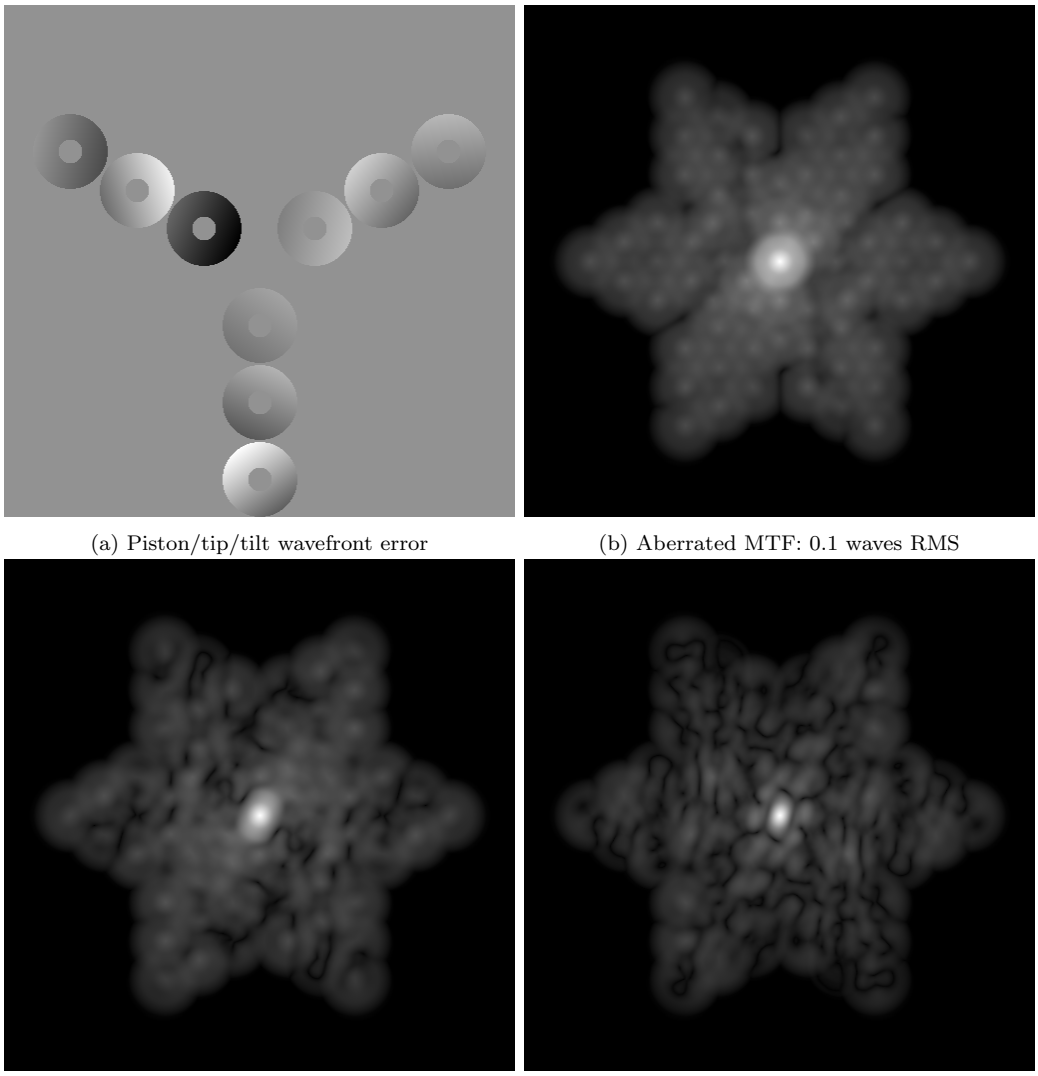
Given the assumption that the aperture mask and wavefront error can be expressed as two-dimensional quantities, the complex pupil function can be expressed as

$$p[x, y] = |p[x, y]|e^{2\pi i W(x,y)} \tag{4.1}$$

where $|p[x, y]|$ is the aperture mask and $W(x, y)$ is the wavefront error, in waves. Note that there is an implicit spectral dependence here, as the waves unit on $W(x, y)$ is given in terms of some reference wavelength, λ_0. The wavelength dependence is made explicit below,

$$p[x, y, \lambda; \lambda_0] = |p[x, y]|e^{2\pi i W(x,y)\frac{\lambda_0}{\lambda}} \tag{4.2}$$

This form of the equation explicitly shows that wavefront error becomes less of an issue as wavelength increases, although diffraction increases to work against this effect. Once this complex pupil function has been constructed for an isoplanatic region in the image, it can be used to predict the image degradation based on the optics through the use of Equation 3.15.



(a) Piston/tip/tilt wavefront error

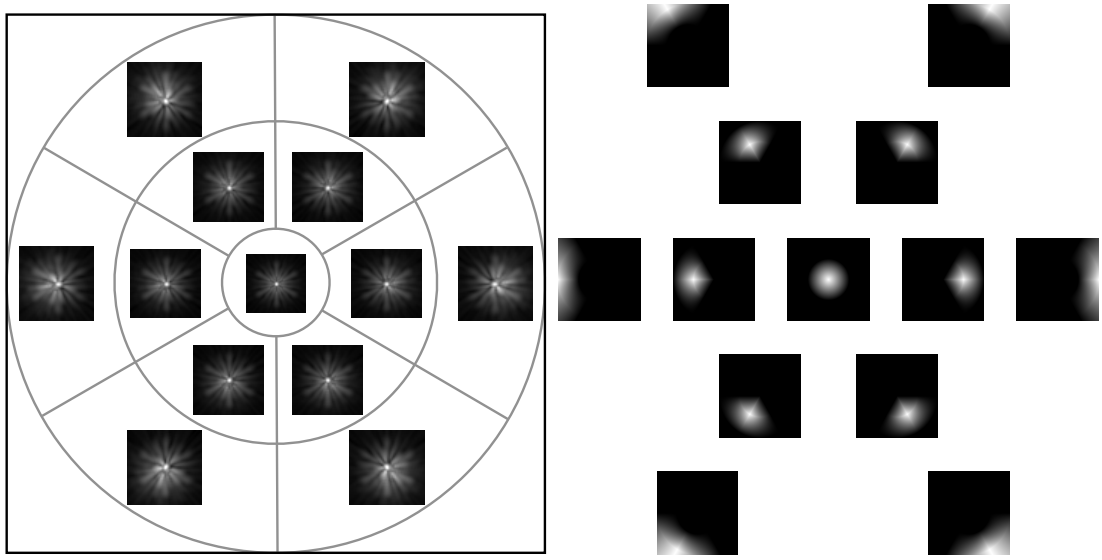
(b) Aberrated MTF: 0.1 waves RMS

(c) Aberrated MTF: 0.25 waves RMS

(d) Aberrated MTF: 0.5 waves RMS

Figure 4.5: Illustration of the effects of wavefront error (piston, tip, tilt) on the MTF of a Tri-arm 9 sparse aperture design.

Figure 4.5 shows an example of how wavefront error can degrade the OTF of a sparse aperture configuration. As can be seen, the OTF quickly degrades with wavefront RMS error. Once the wavefront RMS error exceeds a tenth of a wavelength, there is significant degradation in the periphery of the OTF. For a more detailed analysis of how the MTF degrades with wavefront error, the reader is referred to [Introne, 2004].



(a) PSFs are allowed to vary over the field of view

(b) Interpolation weights for each field point in the final image. Weights are shown on a linear 0-1 grayscale.

Figure 4.6: Off-axis aberrations are modeled by degrading with OTFs computed at several field points and interpolating the results into one final degraded image.

While Figure 4.5 showed how aberrations, such as piston, tip or tilt can degrade OTF performance, these aberrations are invariant with respect to incoming ray angle. Off-axis aberrations, such as coma or astigmatism, do not share this invariance and as such, vary over the field of view of the system. As was previously mentioned, in this modeling approach, wavefront error was approximated as a constant within an isoplanatic region. In order to account for this, the model has the ability to adapt the aberrated optics OTF over the field of view. As is shown in Figure 4.6, the image is split up into a number of radial and angular zones. An OTF is then computed in each zone and used to degrade the image.

After the rest of the degradation pipeline is performed, the degraded images from each isoplanatic region are then used to approximate the final image through interpolation. Bilinear interpolation in polar space is used to derive the interpolation weights. An example of the interpolation weights are shown in Figure 4.6 (b). With this approach, the number of angular and radial zones are user-defined parameters that trade off accuracy with run-time. In this research, 3 radial and 6 angular zones were used.

4.4 Detector Sampling OTF

Unlike a theoretical mathematical sampling, modern imaging detectors are constructed to integrate signal over a finite area. As such, the recorded pixel value is representative of an averaging across the pixel's footprint on the object. In the modeling approach for this research, it will be assumed that pixels are rectangular and have a constant responsivity across their light sensitive areas. With these assumptions, the pixel sampling process can be modeled as (ignoring scale factors),

$$S_{\text{det}}[x,y] \propto \left(RECT\left[\frac{x}{x_p},\frac{y}{y_p}\right] * E_{\text{det}}[x,y] \right) \cdot COMB\left[\frac{x}{p_x},\frac{y}{p_y}\right] \tag{4.3}$$

where, x_p and y_p are the dimensions of the light-sensitive pixel regions and p_x and p_y are the pixel pitches in both dimensions. This equation assumes due to the nature of the $COMB$ function that the light sensitive area is centered in the pixel. This is not a problematic assumption in this research, however, as it will be assumed that $p_x = p_y = x_p = y_p$, meaning that the pixels are square and have a fill factor of one.

With the modeling approach and assumptions just given, the effect of detector sampling can be summarized in an optical transfer function (OTF). Given Equation 4.3, the degradation to the signal is given as a convolution with a $RECT$ function, representing the pixel averaging. As such, the corresponding OTF is given by,

$$OTF_{\text{det}}[\xi,\eta] = SINC[x_p\xi, y_p\eta] = \frac{\sin(\pi x_p \xi)}{\pi x_p \xi}\frac{\sin(\pi y_p \eta)}{\pi y_p \eta} \tag{4.4}$$

An example of the detector sampling OTF is given in Figure 4.7. As can be seen, the effect of detector sampling is relatively minor, when compared to the effect of the optics. In this example, the optical Q of the system was equal to 2. The quantity Q is used in remote sensing to describe the relation between diffraction-limited resolution and detector resolution of a system [Fiete, 1999]. It is given by

$$Q = \frac{\lambda \cdot F\#}{p} \tag{4.5}$$

As can be seen from the figure, the case of $Q = 2$ means that the limiting resolution of the diffraction-limited optics matches with the Nyquist sampling rate of the detector.

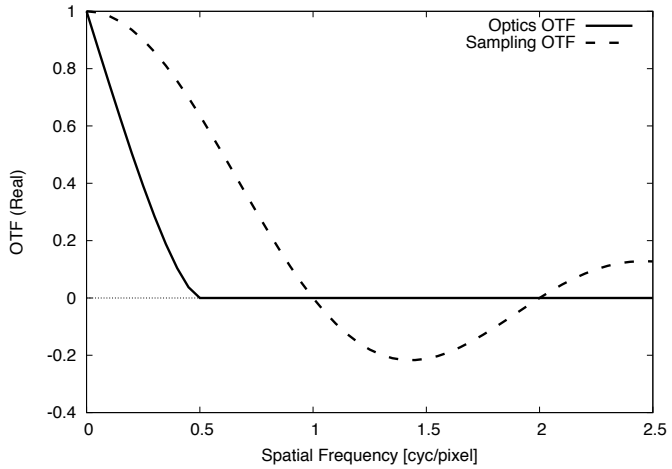


Figure 4.7: Detector sampling OTF in relation to the diffraction-limited circular aperture for a $Q = 2$ system.

Applying the detector sampling OTF makes sense if the model is using radiance data of a higher resolution than that of the modeled detector. However, the normal input to the model is a DIRSIG-generated synthetic image at the detector resolution. As described in the DIRSIG manual, multiple samples are used in determining the radiance reaching each pixel in the system. To do this, multiple strategies may be used. A regular grid of samples in each pixel may be taken and then averaged to give the pixel's radiance value. Another strategy used is adaptive sampling, which utilizes a number of samples randomly cast into a set of sub-pixels. Adaptive sampling then monitors how the radiance estimate converges and terminates the sampling process once the estimate has met some convergence condition. Both of these strategies aim to simulate the detector footprint, which is the same thing the detector sampling OTF is doing to the continuous light distribution on the sensor. As such, it does not make sense to apply the detector sampling OTF in the case of synthetic imagery generated at the modeled detector's resolution, as that would be applying the averaging effect twice.

4.5 Detector Jitter OTF

The previous section discussed how averaging occurred within a pixel's footprint on the ground. However, imaging systems have a finite integration time and over that time the pixel's footprint can move. In order to apply the effects of jitter as an optical transfer

function, it must be a linear and shift invariant effect. This can be justified by assuming that the detector itself is rigid and jitter comes from the platform that the imaging system is attached to, such as a satellite or airplane. As such, the jitter is constant across the detector and the resulting blur can be described as a point spread function.

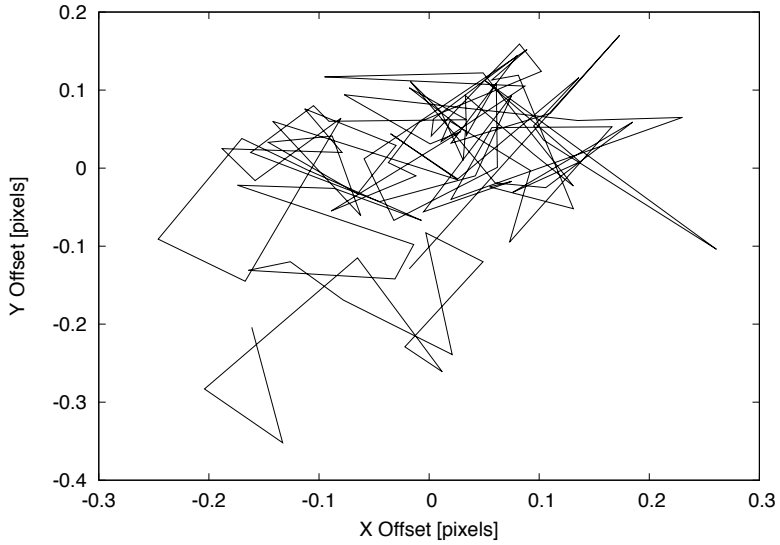


Figure 4.8: An example of a jitter path ($\sigma_j = 0.1$ pixels)

The jitter of any specific platform can vary greatly. For instance, different engines on airplanes can result in different power spectrums of vibration. Vibrations from other instruments on the same satellite can also be potential jitter sources. Precise jitter modeling for a given platform is not the focus of this research, so the model developed simply generated a jitter pattern from a power spectrum. If a specific platform was of interest, its vibration power spectrum could be inserted to get more precise results.

The approach taken in this research to model blur due to jitter was to subdivide the integration time into discrete timesteps and determine the offset due to jitter at each timestep. Given a power spectrum $|J(f)|^2$ for jitter, the square root was taken to find the magnitude of jitter at each frequency and a random phase can be attached. The inverse Fourier transform will then give an offset for each timestep. In this research, all optical components are rigidly mounted to the same optical table, so jitter will be small

and approximately Gaussian. As a result, random numbers from a Gaussian distribution were used for the phase and $1/f$ was used as the power spectrum. The resulting offset pattern was scaled in each dimension to match a user-provided standard deviation σ_j, in pixel dimensions. An example instantiation of a jitter pattern is given in Figure 4.8.

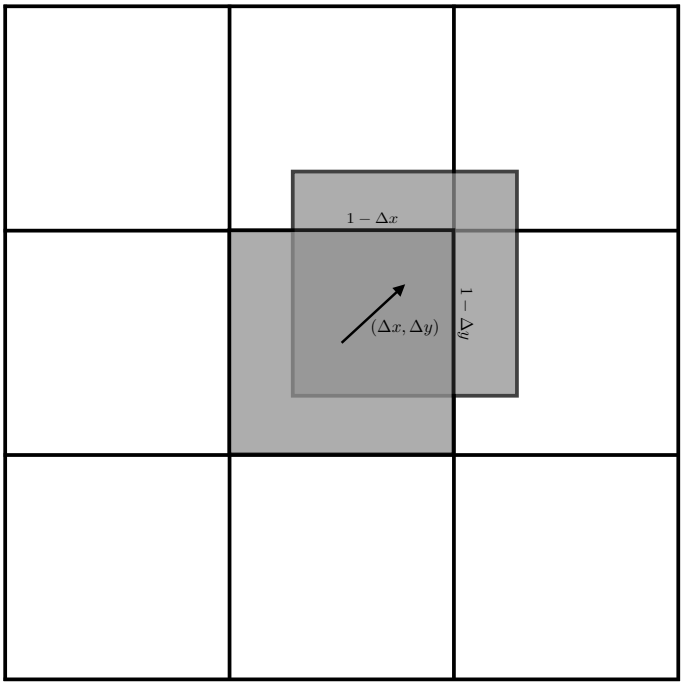


Figure 4.9: Overlap between a displaced pixel due to jitter and the original pixel

In order to translate the generated jitter path into an OTF, a point spread function must be generated. If the pixel's footprint moves by an offset of (Δx, Δ y), as shown in Figure 4.9, than the normalized overlap area between the pixel's new footprint and its old footprint is given by $(1 - \Delta x)(1 - \Delta y)$. The remainder of the pixel's footprint is split between the three neighboring pixels in the original detector space. This example is for pixel shifts of less than one pixel, however, the same idea holds for larger shifts, in that the overlapping area needs to be computed between the shifted pixel and the four pixels in the original grid that the shifted pixel now overlaps. These four overlapping areas now form the point spread function for the corresponding time step in the integration time. If this procedure is repeated for each timestep and the resulting point spread functions are averaged together, an aggregate point spread function for jitter can be derived. The Fourier

transform of this will give the jitter OTF that can be applied as an image degradation in the model. Examples of the resulting MTFs are given in Figure 4.10.

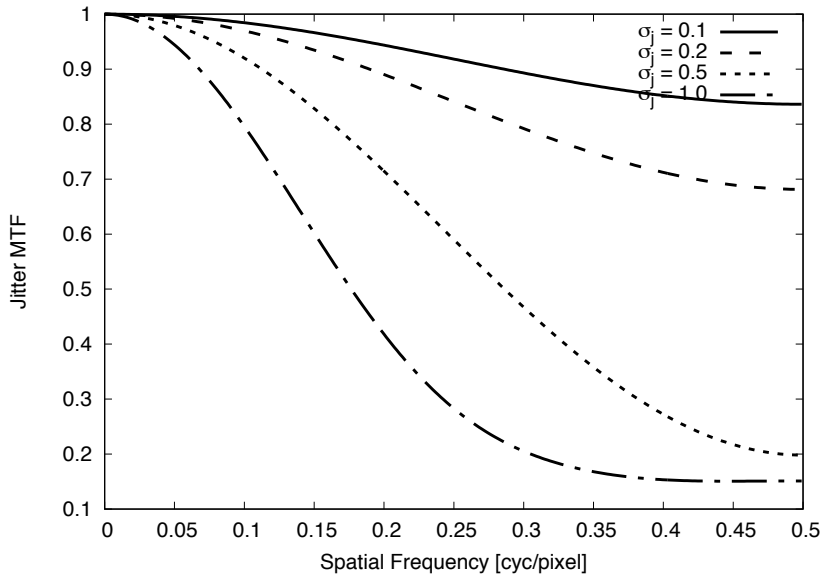


Figure 4.10: Examples of MTFs generated by the jitter modeling procedure. σ_j values are in units of pixels.

4.6 Smear OTF

The previous section described how blur is introduced due to random jitter of the imaging platform. In addition to jitter, there is also deterministic image motion on most remote sensing systems. Both aerial and satellite based systems need to maintain a velocity in order to maintain flight and orbit, respectively. As such, due to the finite integration time of image collection, the detector footprints will be moving linearly on the ground, thus blurring in the direction of flight. This effect is termed smear.

In this research, the imaging platform will be assumed to have a constant altitude. During the image capture process, the platform will be assumed to move with a constant velocity, (v_x, v_y). The speed is then given by $v = \sqrt{v_x^2 + v_y^2}$. Over the integration time, t_{int}, the sensor will move a distance of $t_{\text{int}}v$. Assuming the constant velocity, the blur can be modeled as a $RECT$ function of that width oriented in the direction of motion. The

OTF due to smear is given by the Fourier transform of this blur, given as,

$$OTF_{\text{smear}} = SINC[vt_{\text{int}}\xi] \tag{4.6}$$

In this case, the ξ frequency is still spatial frequency, however, it is frequency in the direction of motion, and not the x-direction, as has been the case in the rest of this document.

Some examples of smear OTFs are given in Figure 4.11. In this example, the integration time was 25 microseconds, the pixel pitch was 7.5 microns and the velocity varied. The velocities were 1-dimensional in the y-direction, so the displacements are computed and given in the legend as Δy values in units of pixels.

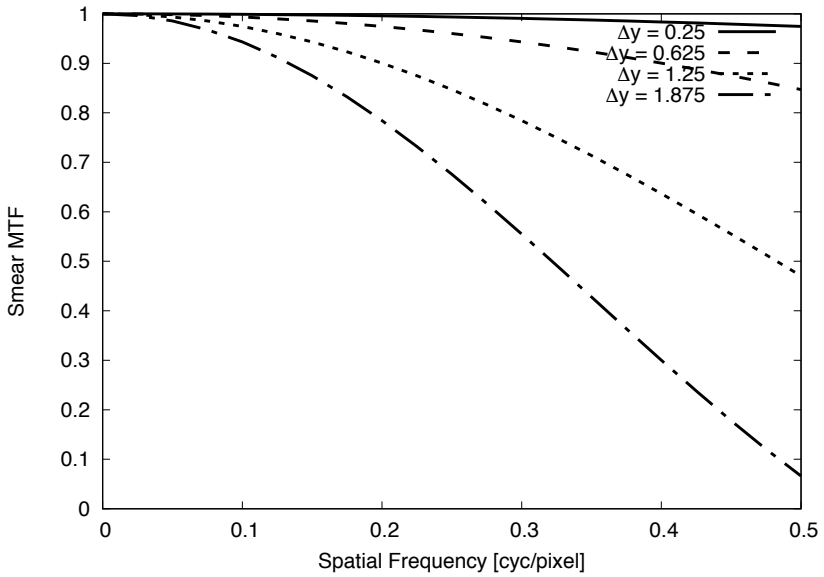


Figure 4.11: Examples of MTFs generated by image smear. Δy values are in units of pixels.

In addition to smear in the primary track direction, cross-track smear can also be introduced by platform motion. Aircraft roll and active pointing systems are two prime examples of mechanisms that can introduce cross-track smear. For aircraft roll, if the velocity of motion in this dimension is available, the image degradation can be approximated with the same MTF expression derived for the primary track direction. However,

image motion due to cross-track roll is technically non-linear, as it changes the angle of projection for each pixel onto the ground. Since this is not present in the laboratory setup in this research, it will not be addressed here. Active pointing systems can be used on framing systems to mitigate smear, but create non-linear smear that varies across the field-of-view. While these systems would be useful for sparse aperture framing systems, the blur introduced is not unique to sparse aperture systems. So, existing smear models for active pointing systems may be applied to sparse aperture systems without any issues and only the linear smear model will be used here for simplicity's sake.

4.7 Detected Image

After applying all of the processing just described, an accurate estimate of the degraded spectral radiance distribution, or L_{source} in the numerator of Equation 3.21, has been computed. This radiance distribution now needs to be converted into detector response. In order to perform this conversion, the user needs to provide a specification for the detector being used. The specification is described in Table 4.1.

Table 4.1: User-provided parameters for the modeled detector

Parameter	Description	Unit
$QE_b(\lambda)$	Quantum efficiency spectrum for each band	unitless
p	Spacing between detector pixels	meters
F_{det}	Fill factor of the detector pixels	unitless
T	Operating temperature of the detector	Kelvin
T_d	Dark current doubling temperature	Kelvin
T_{ref}	Dark current reference temperature	Kelvin
$\sigma_{\text{dc,ref}}$	Dark current reference RMS	RMS e^- / pixel / s
$N_{\text{full-well}}$	Full-well capacity of each pixel	e^-
n	Bit depth of the analog-to-digital converter	bits
S_{ADC}	"Effective" gain factor of the detector	e^- / digital count
σ_r	Read noise RMS	RMS e^-
t_r	Readout time	seconds

From the specifications given for the detector, the area of the detector element is computed as

$$A_{\text{det}} = p^2 F_{\text{det}} \tag{4.7}$$

This area is used to convert the irradiance distribution on the detector, produced by Equation 3.21, to a spectral flux distribution in photons.

$$\Phi_p(x, y, \lambda_i) = A_{\text{det}} \cdot t_{\text{int}} \cdot E_{\text{det}}(x, y, \lambda_i) \cdot \frac{\lambda}{hc} \tag{4.8}$$

If this photon flux estimate was generated using synthetic data from DIRSIG, then photon noise has not yet been included in the estimate and must be added now. Photon noise is characterized by a Poisson distribution, where the mean and variance are the expected signal level, Φ_p. For reasonably large signal levels (approximately 10 photons), the Poisson distribution can be well approximated by a normal distribution, where the standard deviation is the square root of the signal level. Thus, the signal with photon noise included is given by

$$\tilde{\Phi}_p(x, y, \lambda_i) = \Phi_p(x, y, \lambda_i) + \mathcal{N}\left(0, \sqrt{\Phi_p(x, y, \lambda_i)}\right) \tag{4.9}$$

The signal produced by the detector in electrons, S_e, for a given band, b, is then given by

$$S_e(x, y, b) = \sum_i \tilde{\Phi}_p(x, y, \lambda_i) \cdot QE_b(\lambda_i) \tag{4.10}$$

where $QE(\lambda_j)$ is the linearly interpolated quantum efficiency, calculated as

$$QE_b(\lambda) = \left(\frac{\lambda - \lambda_{i-1}}{\lambda_i - \lambda_{i-1}}\right) QE_b(\lambda_i) + \left(1 - \frac{\lambda - \lambda_{i-1}}{\lambda_i - \lambda_{i-1}}\right) QE_b(\lambda_{i-1}) \tag{4.11}$$

where λ_i is wavelength of the QE sample that is closest to the wavelength of interest, λ, while still being greater than λ. As such, λ_{i-1} is the wavelength of the QE sample that is closest to λ while also being less than λ.

The detector noise is then added to this signal in electrons. The detector noise is treated as a sum of independent noise sources for the read noise and the dark current noise. The read noise RMS is given by the user in electrons and is added as a Gaussian noise source directly to the signal. For a given detector temperature, T, the dark noise RMS is computed as

$$\sigma_{\text{dc}} = t_{\text{int}} \cdot 2^{\frac{T - T_{\text{ref}}}{T_d}} \cdot \sigma_{\text{dc,ref}} \tag{4.12}$$

where T_d is the dark current doubling temperature and T_{ref} is the reference temperature

for $\sigma_{\text{dc,ref}}$. The dark current noise is then computed as a Gaussian noise source in electrons and also added to the signal. The user-provided analog-to-digital conversion gain, S_{ADC}, and quantization bits, n, can then be used to quantize the signal into digital counts by the equation,

$$I(x, y, b) = \min\left(2^n, \left\lfloor \frac{S_e(x, y, b)}{S_{\text{ADC}}} \cdot \right\rfloor\right) \tag{4.13}$$

where $\lfloor \cdot \rfloor$ is the integer "floor", or round-down, operation. This equation assumes the detector is linear over the whole range of light levels up to its full-well capacity. A model for how the detector behaves approaching saturation could be integrated with little trouble.

4.8 Restoration Filter

As was seen in the previous sections on the OTF of the system optics, sparse aperture systems have transfer functions that are non-zero for a large range of frequencies, however, the magnitude of the transfer function at those frequencies is low. As such, inverse filtering is a critical step in restoring image quality and extracting information from these images. As was shown in Section 3.7, the inverse filter that will be used in this research is given by

$$\hat{F}[\xi, \eta] = W[\xi, \eta] \cdot G[\xi, \eta] = \frac{H^*[\xi, \eta]}{|H[\xi, \eta]|^2 + \gamma \cdot |S[\xi, \eta]|^2} \tag{4.14}$$

In this equation H is the optical transfer function of the system and S is the transfer function of some smoothness operator. In this research, the smoothness operator was a 3×3 Laplacian convolution kernel, shown in Figure 4.12.

0	-1	0
-1	4	-1
0	-1	0

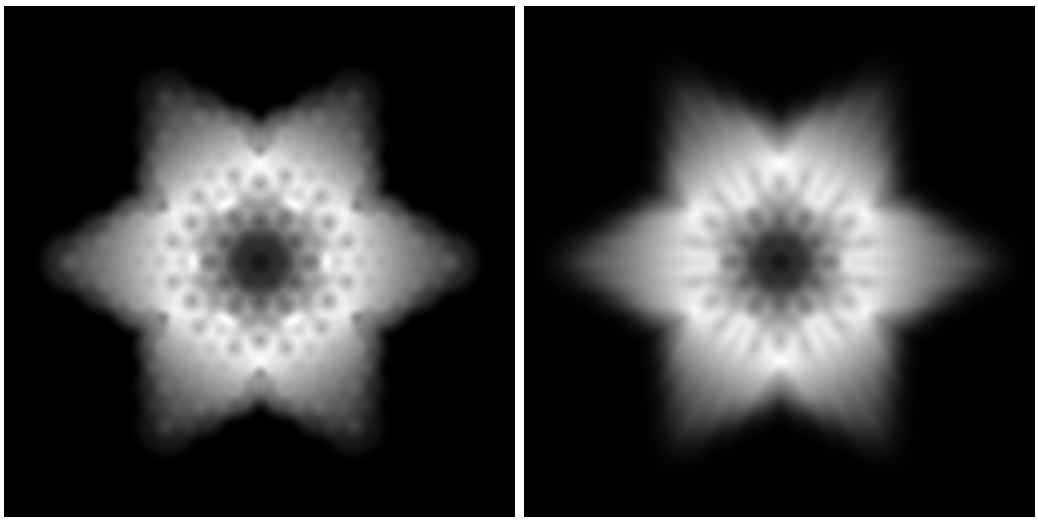
Figure 4.12: 3×3 Laplacian convolution kernel

The transfer function, on the other hand, is less straightforward. Since this inverse filter is applied to a signal integrated over a wide bandpass, it does not make sense to use the system OTF at any specific wavelength. Instead, an "effective" OTF must be constructed for the entire bandpass. In this research, the OTFs for wavelengths over the bandpass were averaged in accordance with the relative spectral responsivity curve of the

band. In addition, if the illumination spectrum of the scene is known, it can also be used to guide the weighted average of the effective OTF.

$$OTF_{\text{eff,b}} = \frac{\sum_i OTF(\lambda_i) \cdot R(\lambda)}{\sum_i R(\lambda_i)} \tag{4.15}$$

For sparse aperture system that have OTFs with many small peaks, this averaging effectively blur those peaks radially in the frequency domain. This effect is minor for typical hyper spectral bandwidths, however, it is drastic for panchromatic bands. The averaging effect was shown in Figure 4.4. Since Equation 4.14 depends on the effective OTF, the effects of spectral averaging are also seen in the restoration filter. The restoration filters for a Tri-arm 9 configuration with a 10 nanometer FWHM band and a 100 nanometer FWHM band are shown in Figure 4.13. The peripheral peaks of the Tri-arm 9 OTF are still clearly present in the 10-nanometer bandpass but have been blurred away in the 100 nanometer bandpass.

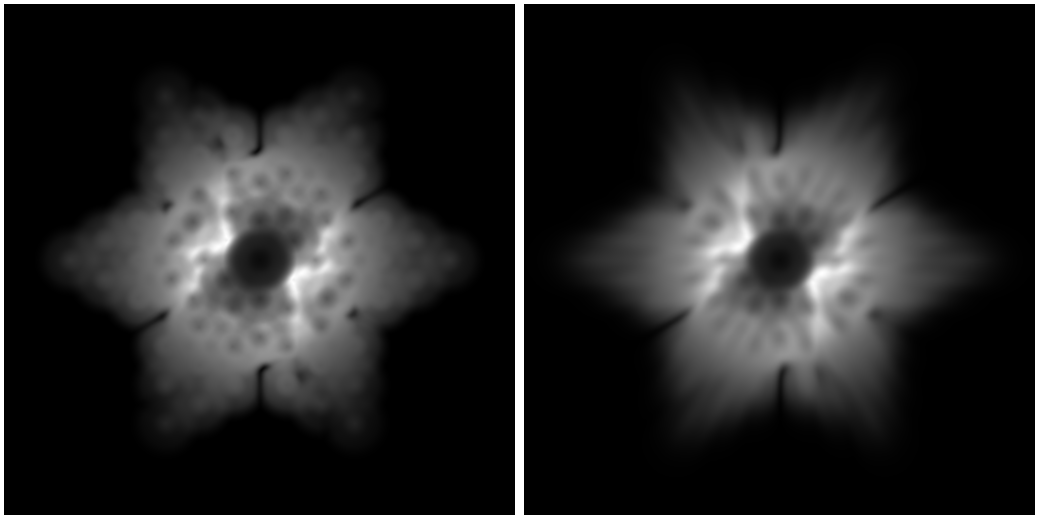


(a) Center: 550 [nm], FWHM: 10 [nm] (b) Center: 550 [nm], FWHM: 100 [nm]

Figure 4.13: Illustration of the effects of the spectrally averaged OTF on the restpratopm filter.

The restoration filters shown in Figure 4.13 work very well when applied to unaberrated imagery degraded with the described model. However, wavefront error can quickly degrade the OTF of a sparse aperture system, as was shown in Figure 4.5. Wavefront error

has the effect of eliminating some of the peaks in the OTF. If the inverse filters shown in Figure 4.13 are used to restore images that were imaged with wavefront error, then frequencies not present in the degraded imagery will be boosted, leading to undesirable artifacts that degrade image quality. This is a large issue in a realistic imaging scenario, as an estimate of the wavefront error may be unavailable or may require an advanced technique, such as phase retrieval, to obtain. [Miller et al., 2007] showed that while phase retrieval can obtain wavefront error estimates from sparse aperture imagery, it is only effective under constrained circumstances, which may be difficult to reproduce in a realistic system.



(a) Center: 550 [nm], FWHM: 10 [nm] (b) Center: 550 [nm], FWHM: 100 [nm]

Figure 4.14: Illustration of the effects of wavefront error (0.1 RMS waves of piston, tip, tilt) on the inverse filter.

If however, a wavefront error estimate can be obtained, it can incorporated into the inverse filter. Equations 4.14 and 4.15 already handle the OTF produced by the complex aperture function, so simply using the estimate of the wavefront error to predict the OTF will result in a compensated inverse filter. Figure 4.14 shows inverse filters under the same conditions as Figure 4.13 with the addition of wavefront error. In this case, random piston, tip and tilt error were added so that the RMS wavefront error was 0.1 wavelength or 55 nanometers, in this case. As can be seen, there are specific peaks that have been

suppressed in the inverse filter. However, there are also regions closer to the origin that are boosted more. These are regions of low but non-zero support in the aberrated optics OTF. The high boosting of these regions can recover resolution but can also create undesirable artifacts.

While the Wiener filtering approach will be used in this study, it should be noted that there are many other deconvolution algorithms that may also perform well on sparse aperture imagery. [Fienup et al., 2002] performed a study that compared Wiener filtering to a maximum likelihood reconstruction algorithm and found that the two algorithms performed equivalently for reasonable SNR values. Finding the ideal post processing algorithm is not the focus of this research, so the accepted standard of the Wiener filter will be used.

Chapter 5

Laboratory Model Validation

5.1 Introduction

The first key objective of this research will be to validate the sparse aperture image quality model, set forth in Chapter 4. This model has been implemented using C++ and runs on any Linux computer with a modern (C++11) compiler and the appropriate dependencies. One of the key features of the model is that the user has direct control over the input radiance field, detector and the aperture configuration that are used in the modeled system. Given an imaging detector, the user can recreate that detector in the model by entering values commonly found in the detector's specifications. As for apertures, a number of common sparse aperture components are built into the model. The user can then create their own apertures through a configuration file, where they can specify which sub-apertures to use and their positions, scales and rotations, if necessary. If a component of the aperture is not included in the model, a simple programming interface exists to add it. Wavefront error is also fully supported, through Zernike polynomials or an input data file. These controls allow for the recreation of a small-scale laboratory system in the computer model, a prerequisite for a validation study.

In previous work, the modeling approach from Chapter 4 was shown to predict noticeable artifacts in imagery produced after inverse filtering. An example of this is shown in Figure 5.1. As can be seen, ringing artifacts can arise after post-processing on sparse aperture imagery. Post-processing of traditional imagery results in simple edge-overshoot artifacts, but more complex artifacts arise due to the more complicated OTFs of sparse

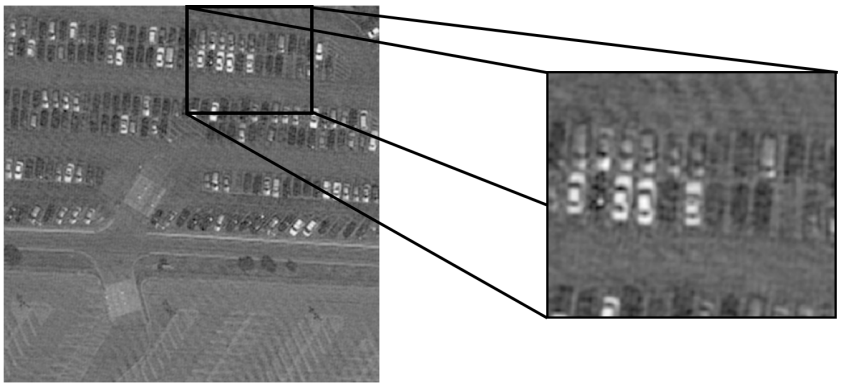


Figure 5.1: Predicted ringing artifacts due inverse filtering over large bandpasses [Block, 2005]. The ringing is most noticeable in the pavement areas.

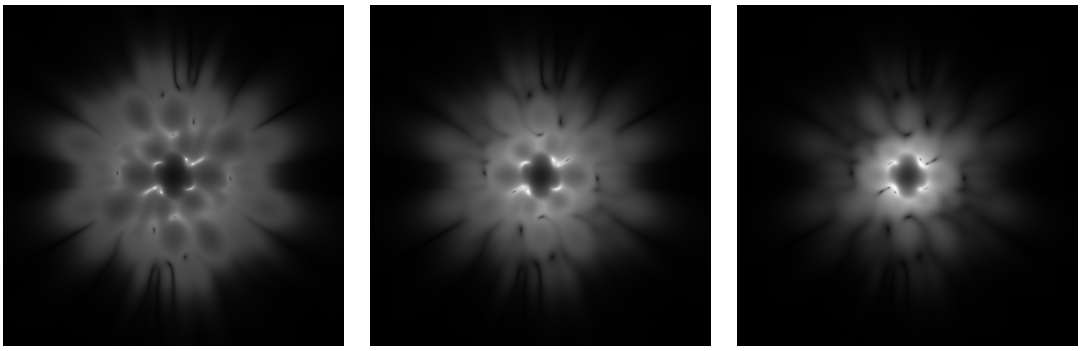
aperture systems. While some artifacts are inherent in post-processing, they are exacerbated by errors in the post-processing filter. As was shown in Section 4.8, the post-processing filter is given as,

$$\hat{F}(\xi,\eta) = W(\xi,\eta)\cdot G(\xi,\eta) = \frac{H^*(\xi,\eta)}{|H(\xi,\eta)|^2 + \gamma\cdot|S(\xi,\eta)|^2}\cdot G(\xi,\eta) \tag{5.1}$$

Examples of some post-processing filters are given in Figure 5.2, showing the effect of the smoothness term in the denominator. As can be seen in the figure, the filters have a similar level of detail to the sparse aperture OTF and thus require significant spectral sampling to model correctly. Since H is the only unknown in the post-processing filter, any errors in the post-processing filter derive soley from errors in the approximation of the effective OTF. Over a bandpass, the effective OTF of the system is given as

$$H_{\text{eff}}(\xi,\eta) = \frac{\int_0^\infty SW(\lambda)\cdot H(\xi,\eta,\lambda)\ d\lambda}{\int_0^\infty SW(\lambda)\ d\lambda} \tag{5.2}$$

where, $H(\xi,\eta,\lambda)$ is the monochromatic OTF of the imaging system and $SW(\lambda)$ is the spectral weighting function over the system's bandpass. Errors can arise in both terms. Errors in the monochromatic OTF normally arise from errors in the wavefront error estimates. Errors in the spectral spectral weighting function normally arise from incomplete



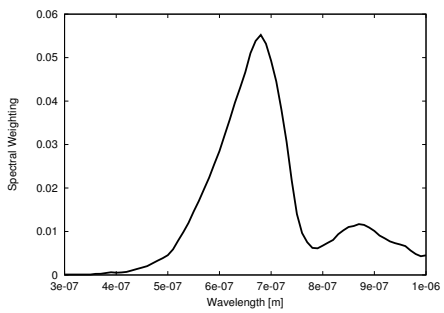
(a) $\gamma = 1 \times 10^{-4}$ (b) $\gamma = 1 \times 10^{-3}$ (c) $\gamma = 1 \times 10^{-2}$

Figure 5.2: Examples of Golay-6 inverse filters for an aberrated system. γ is the Lagrange multiplier for the smoothness term in Equation 4.14 and modulates how much the secondary peaks are boosted.

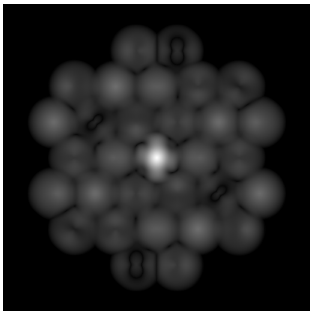
or imperfect knowledge of the scene or the sensor. In a scene-independent analysis, illumination is unknown and the spectral weighting function would be the product of the optics transmission spectrum, the detector responsivity over the bandpass and at best a nominal scene spectrum. In a laboratory setting, the light source spectrum can also be included.

Since the spectral scaling in the system OTF (Equation 3.15) is uniform with respect to ξ and η, the integration over the bandpass has the effect of radially averaging the OTF. This effect is illustrated in Figure 5.3. As can be seen in the figure, the spectral averaging has the effect of "filling in" the gaps between peaks. However, in the presence of wavefront error, it also has the effect of spreading out the holes in the OTF. These extended valleys in the MTF contribute to artifacting, especially if not accounted for in the post-processing filter. [Block, 2005]

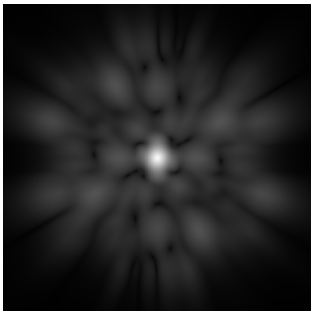
This chapter will describe an experiment to validate the OTF predictions of the model described in Chapter 4. The laboratory setup will be described, with detailed descriptions of each component. The system's MTF will be directly measured and compared to the model predictions. Post-processing artifacts will also be observed, but in-depth discussion can be found in Chapter 6.



(a) Spectral weighting



(b) Monochromatic MTF (Gamma Scale)



(c) Effective MTF (Gamma Scale)

Figure 5.3: An example of a spectrally-averaged MTF for a Golay-6 sparse aperture. The spectral weighting function is the product of tungsten source spectrum with an IR-reject filter and a silicon CCD quantum efficiency.

5.2 Laboratory System

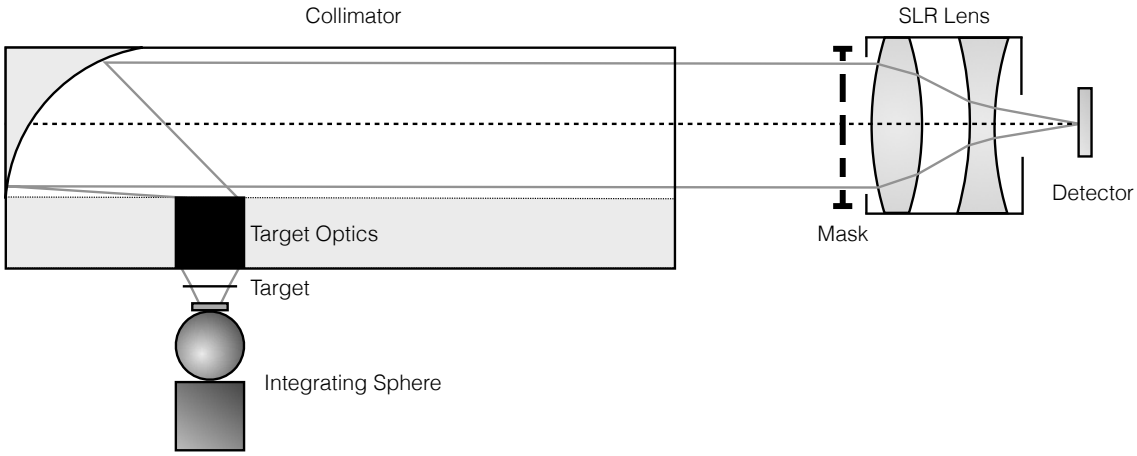


Figure 5.4: Schematic of laboratory validation study optical setup.

Given the modeling approach described in Chapter 4, the purpose of this research is to validate the predicted results in a laboratory setting. This presents the challenge of constructing a sparse aperture system. This task has been attempted before, namely by [Chung et al., 2002] and [Zhou et al., 2009]. Chung et. al. took the approach of a direct construction, however, the cost was hundreds of thousands of dollars, putting it out of

the budget for this research. Zhou et. al. took the approach of a $4f$ optical system with a sparse aperture mask. This research draws more heavily from the latter approach, however, in an attempt to perform broadband experiments and introduce wavefront error into the system, the design will be modified slightly. A schematic of the design is given in Figure 5.4.

Each component of the system will be described in the following sections. The key insight in this design, however, is that the mask must be the aperture stop of the system. As such, it will dominate the MTF due to diffraction.

5.2.1 Integrating Sphere

The integrating sphere used in this system functions as the light source. Integrating spheres have the desirable property of providing uniform illumination, which will make modeling the source-object interactions easy, especially for binary targets, such as edge targets. The sphere, pictured in Figure 5.5, is a calibrated sphere from Optonics Laboratories. The calibration is unfortunately out-of-date. For the model validation study, however, the calibration is not necessary. Simply characterizing the current spectral output from the sphere will be sufficient for this study.

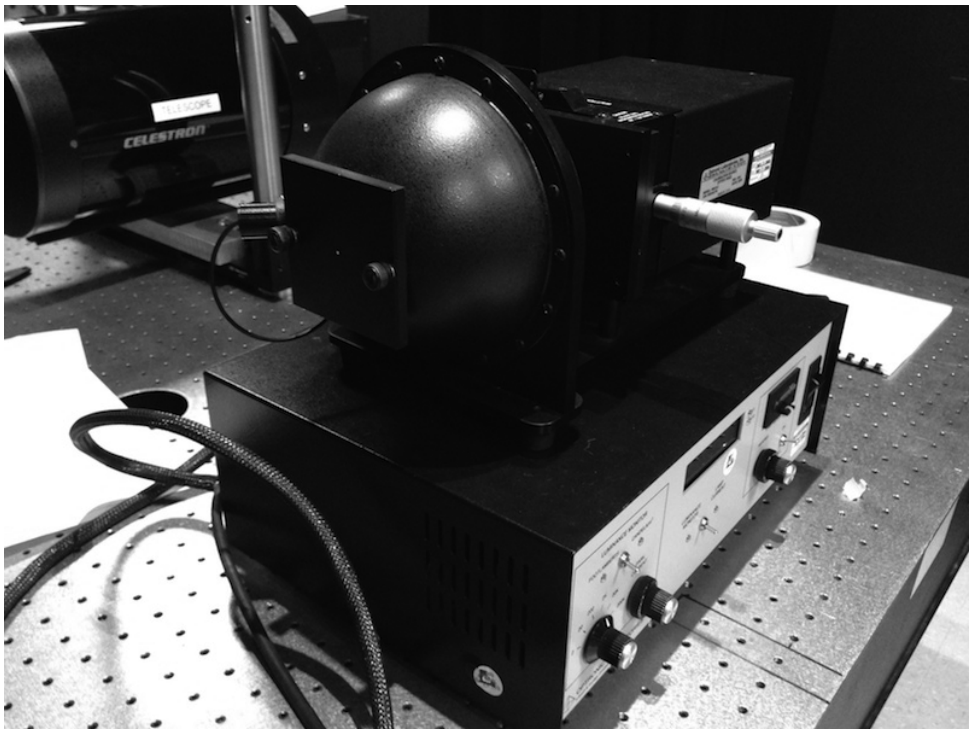


Figure 5.5: Integrating sphere from Optonics Laboratories

As has been established in previous chapters, the spectral dimension is very important to modeling sparse aperture system performance. As such, it was important to characterize the spectral output of the source used in laboratory experiments. To do this, an ASD Inc. spectrometer was used. A 3° field-of-view optic was attached to the instrument's fiberoptic and positioned at the exit pupil of the integrating sphere. One hundred samples were taken and averaged to give the radiance spectra of the light source. When this measurement was performed, the light source was being driven at a current of 5.787 Amps. This is important, as the color temperature of the light output by the sphere varies with current. The aperture setting, on the other hand, is not important as that essentially acts as a scalar multiplier on the shape of the spectrum. The spectrum is plotted in Figure 5.6.

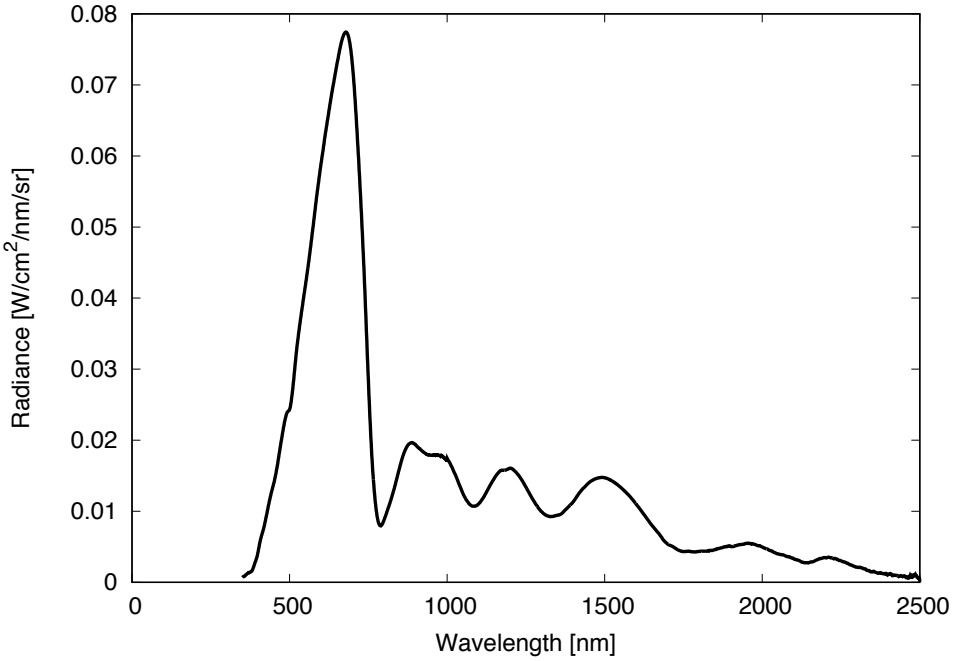


Figure 5.6: Radiance spectrum output from the integrating sphere when driven at 5.787 Amps.

As can be seen from the plot, the peak of the spectral curve is in the visible range, precisely at 679 nanometers. An infrared-reject filter is present in the integrating sphere, which explains why this spectrum is not a closer approximation of a blackbody emission spectrum. This spectrum was also measured with an OceanOptics spectrometer, which produced an identical measurement. The spectral output of the light source will be used heavily in model validation.

5.2.2 Target / Collimator

Since this study focuses on the remote sensing applications of sparse aperture systems, the object plane will be nearly at infinity. To simulate this in a laboratory environment, a collimator needs to be used. The collimator used in this research is an LC-06 from Electro Optical Industries (EOI). This collimator uses an off-axis parabola for its primary mirror. As such, the optical path is free of any obstruction and on-axis aberrations are very low. Wavefront quality is high along the optical axis at 1/8th of a wavelength at 633 nanometers.

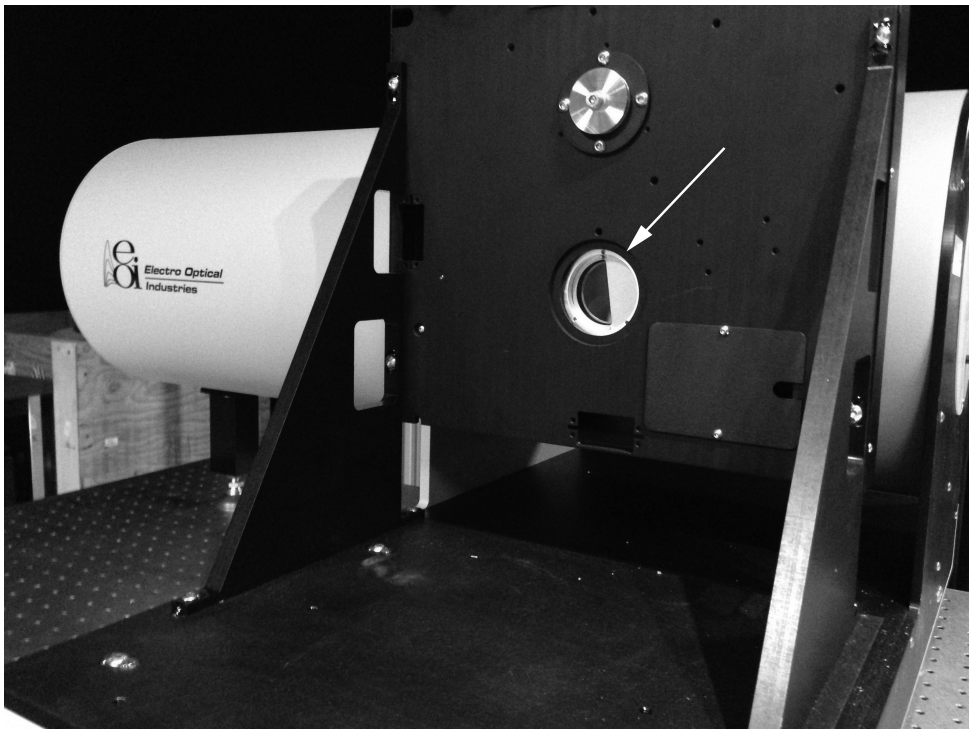


Figure 5.7: Target plane of the EOI LC-06 collimator

The target plane of the collimator is shown in Figure 5.7. As can be seen, an edge target is installed. The target can be freely rotated to any orientation. This edge target has been slightly inclined relative to the axis of the imaging system, so that the slanted-edge MTF measurement technique may be used. The collimator has a multi-position target wheel, so additional targets may be inserted. Targets must fit inside the 2-inch diameter circular holder in the target wheel. While the edge target is sufficient for MTF validation, alternative targets will be needed for other experiments.

Once the target has been placed into the target wheel, it is collimated by the internal optics of the collimator. For the LC-06 model, the effective focal length of the system is 30 inches. The clear aperture of the system is 6 inches, giving an F-number of 5. Another view of the collimator is given in Figure 5.8.

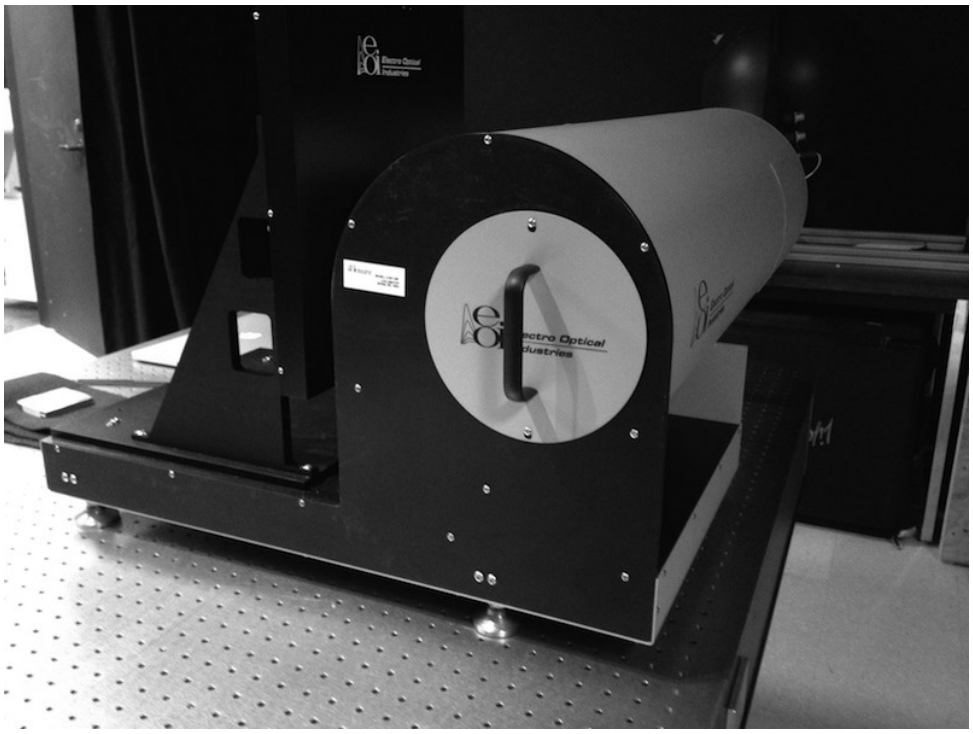


Figure 5.8: EOI LC-06 collimator exit aperture (cover on)

5.2.3 Mask

The mask is the key element of the system and is responsible for simulating the sparse aperture diffraction effect. For this experiment, a mask with the Golay-6 aperture pattern and an 18% fill factor was used. This mask was produced by Applied Image Inc. as a chrome-on-glass print. The accuracy of the print was quoted at 0.1-0.25 microns, well below the size of the printed sub-apertures. While the entire system may bear little resemblance to a sparse aperture telescope, if care is taken to ensure that the mask is the aperture stop of the system, then the diffraction-limited OTFs of the systems will be equivalent. The encircled diameter of the sub-apertures on the mask is 3.4 millimeters. This size was chosen with knowledge of the camera system and resulted in the mask being the aperture stop of the system, with a system $f/\#$ of 24. Two photographs of the laboratory setup and the sparse aperture mask are given in Figure 5.9. As will be shown later, this configuration is well-suited for the validation of aberrated MTF predictions. Through the examples in previous chapters, it has been shown that the MTF of a sparse aper-

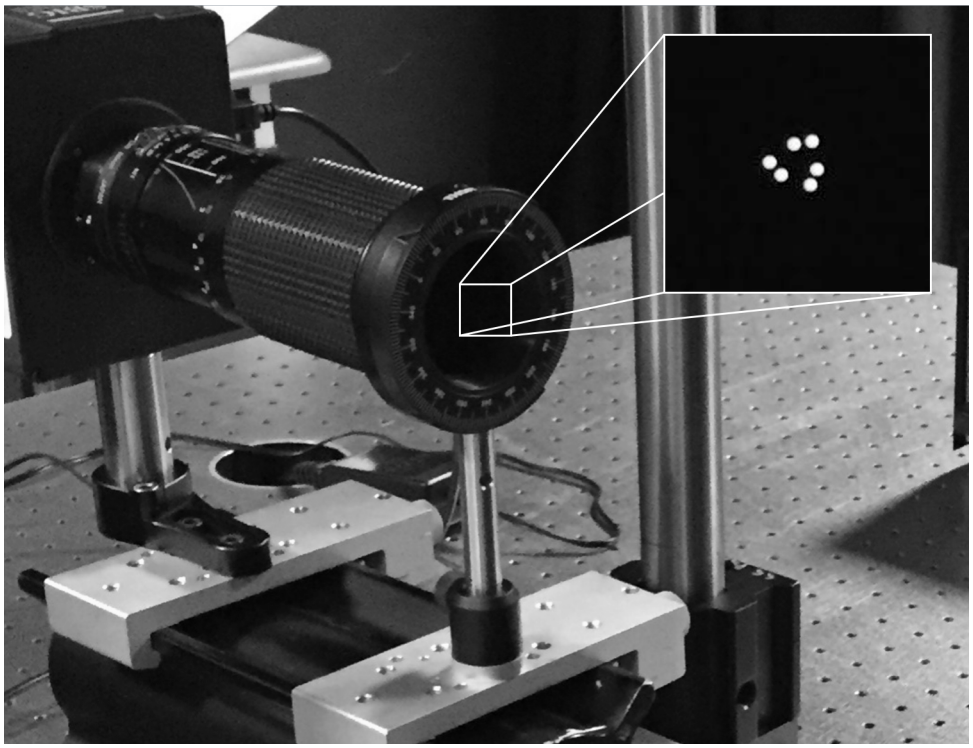


Figure 5.9: Golay-6 sparse aperture mask mounted on the back of a rotation stage.

ture system can be highly angularly-dependent. Since the slant-edge MTF technique only produces a 1-dimensional slice of the MTF, simply taking one measurement will not be sufficient. While the edge target can rotate in the target wheel, the efficiency of the slant edge method is at a maximum when the edge is oriented slightly off-axis. Thus, keeping the edge at a high-efficiency position, while rotating the mask, is a more desirable solution to sweeping out the 2D MTF with a number of 1D radial profiles. As such, the aperture mask was mounted onto the back of a rotational stage, which had 2 degree markings.

5.2.4 Imaging Camera

Once the sparse aperture mask has degraded the input radiance field, an image is formed using a panchromatic imaging system. An Albinar 80-200mm zoom lens is used for focusing. In all experiments in this research, this lens is set at an 80mm focal length. Once

focused, this signal is imaged by an SBIG 8300M astronomical CCD detector. Both of these components are shown in Figure 5.10.

(a) Albinar 80-200mm zoom lens (b) SBIG STF-8300M imaging sensor

Figure 5.10: Components of the imaging portion of the laboratory sparse aperture system

Table 5.1: Relevant detector specifications for SBIG STF-8300M detector

Parameter	Description	Value
p	Spacing between detector pixels	5.4 [μm]
T	Operating temperature of the detector	258.15 [K]
T_d	Dark current doubling temperature	5.8 [K]
T_{ref}	Dark current reference temperature	258.15 [K]
$\sigma_{\text{dc,ref}}$	Dark current reference RMS	0.02 [e^-/pix/s]
$N_{\text{full-well}}$	Full-well capacity of each pixel	25500 [e^-]
n	Bit depth of the analog-to-digital converter	16
S_{ADC}	"Effective" gain factor of the detector	0.37 [e^- / count]
σ_r	Read noise RMS	9.3 [e^-]

As has been mentioned in previous chapters, sparse aperture imaging is notorious for having signal-to-noise concerns. As such, it is critical that the imaging detector has good noise performance. The relevant scalar detector parameters are given in Table 5.1. The quantum efficiency is plotted in Figure 5.11. In all experiments run in this research, the light source was adjusted such that integration time could be set to less than half a second. As such, dark current was not a concern and system noise came from read noise and photon noise.

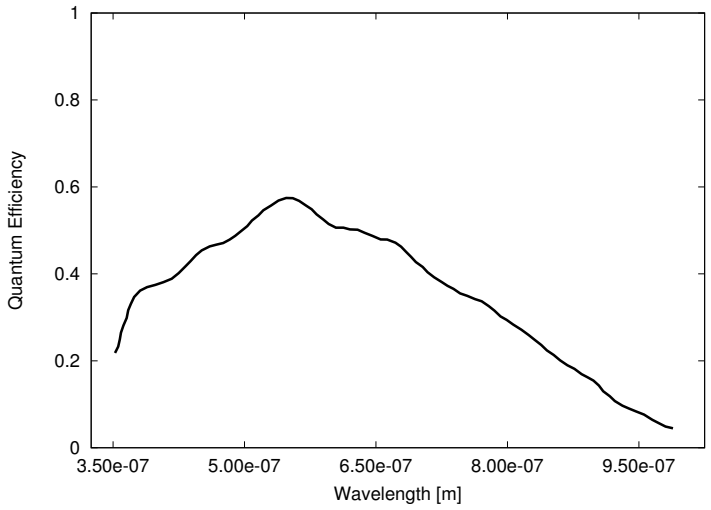
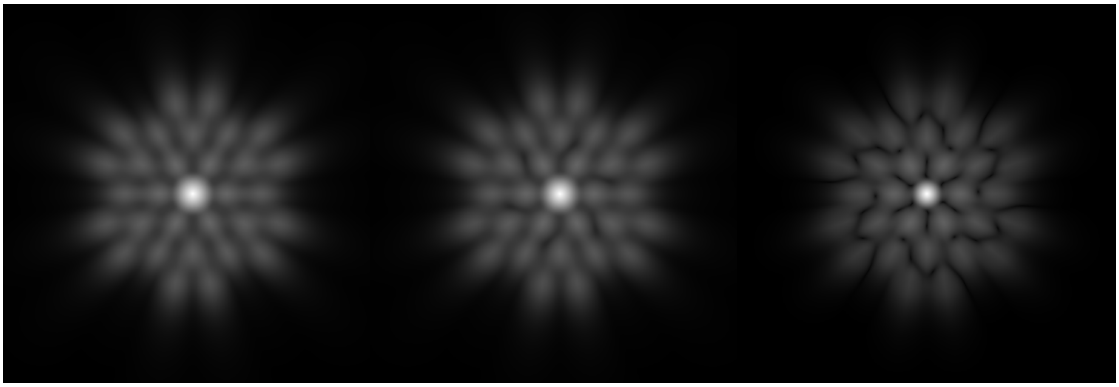


Figure 5.11: SBIG STF-8300M quantum efficiency spectrum

5.3 Methodology

Given the laboratory setup just described, the goal of this experiment is to validate the OTF predictions of the sparse aperture model. In practice, we will approximate the OTF with its complex magnitude, the MTF. In order to obtain MTF measurements from the laboratory system, the slanted edge method, introduced by [Burns, 2000], will be used. As input to this method, a region of interest of roughly 100 pixels was extracted around the center of the edge to perform the analysis. This was done for a number of reasons. Due to the off-axis parabolic design of the collimator, the area of high image quality is limited to a small field angle. In addition, if analysis is confined to the on-axis region, the only lens aberrations of concern are defocus and spherical aberration.

Since post-processing artifacts only arise in the presence of wavefront error, it was also necessary to validate the model's predictions under this error. Due the construction of the system, each individual sub-aperture cannot have independent aberrations. This is unfortunate, as introducing a first-order aberration, such as piston, to only one sub-aperture, results in clear MTF effects that are confined to a subset of the MTF peaks. However, due to the small size of the sub apertures, introducing these errors with additional optics was not possible. As was previously mentioned, by confining analysis to the center of the image, the only higher-order aberrations of concern are defocus and spherical aberration



(a) 0 waves defocus (b) 0.5 waves defocus (c) 1.1 waves defocus

Figure 5.12: Effects of defocus on theoretical Golay-6 MTF. All images are contrast-stretched. Defocus does not effect peak location, but does increase the width of the gaps between peaks.

from the imaging lens. In the range around the mask's $f/\#$ of 24, no spherical aberration was detectable, so defocus could be used to introduce controllable wavefront error into the system. In this experiment, two levels of defocus, 0.5 and 1.1 waves, were measured and compared to the model. Since the falloff slope of the central peak of the MTF increases steadily with defocus and can be measured with a very high accuracy, this was used to determine the amount of defocus in the system.

In order to determine the validity of the model, some comparison must be drawn between the measured and predicted MTF results. Root-mean squared (RMS) error in frequency space is the simplest metric and gives compact representation of the quality of the fit between the model and measurement. However, it would also be desirable to put the results in terms of image quality. The General Image Quality Equation (GIQE) is used to predict image quality on the National Imagery Interpretability Rating Scale (NIIRS) as a function of system parameters [Leachtenauer et al., 1997]. Utilizing GIQE when predicting sparse aperture image quality is problematic, as the equation was not regressed on sparse aperture images, however, it does give some insight into the relative importance of edge response with respect to ground sample distance and signal-to-noise ratio. Additionally, the artifacts predicted by the model due to inverse filtering cannot be characterized by a simple edge overshoot term. However, a simple analysis can be performed by inverse filtering on

the measured images with filters derived from both the modeled and measured OTFs to determine if any difference is noticeable. Post-filtered relative edge response (RER) can be measured on both images and a rough estimate of the change in NIIRS between the modeled and measured systems can be calculated, giving a perceptual difference between the modeled and laboratory systems.

5.4 Control Experiment

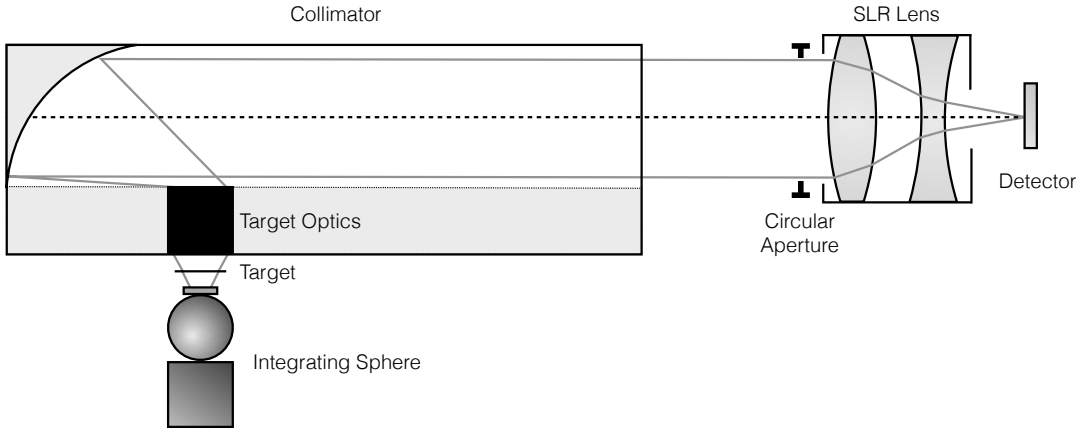


Figure 5.13: Setup for a control experiment to determine the non-modeled MTF of the laboratory system

Since the slanted-edge method was chosen for this experiment, the measured MTFs will be *system* MTFs. That is, they will include the image degradation from every step of the imaging process. As such, a control experiment is required to isolate the effects of the sparse aperture. It is common knowledge that the monochromatic MTF of a circular aperture is,

$$MTF\left(\frac{\xi}{\xi_{\text{cutoff}}}\right) = \frac{2}{\pi}\left(\cos^{-1}\left(\frac{\xi}{\xi_{\text{cutoff}}}\right) - \left(\frac{\xi}{\xi_{\text{cutoff}}}\right)\sqrt{1-\left(\frac{\xi}{\xi_{\text{cutoff}}}\right)^2}\right) \tag{5.3}$$

Over a bandpass, the cutoff frequency varies linearly with wavelength and Equation 5.2 is used to perform spectral averaging. Using this knowledge, a circular aperture was place into the system in lieu of the sparse aperture mask and MTF was measured using the

same slanted-edge procedure. The MTF measured here would then be

$$MTF_{\text{ctrl,meas}} = MTF_{\text{coll}} \cdot MTF_{\text{cir}} \cdot MTF_{\text{lens}} \cdot MTF_{\text{det}} \tag{5.4}$$

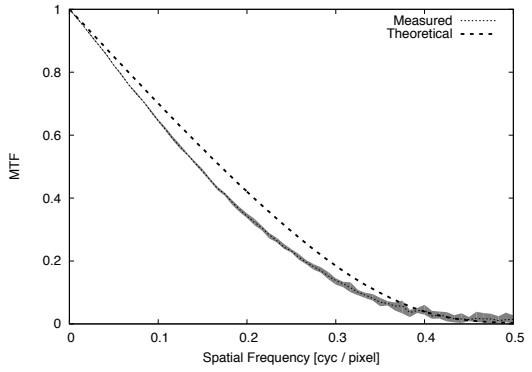
If this system is modeled as described in Chapter 4, the predicted MTF will be

$$MTF_{\text{ctrl,mod}} = MTF_{\text{cir}} \cdot MTF_{\text{det}} \tag{5.5}$$

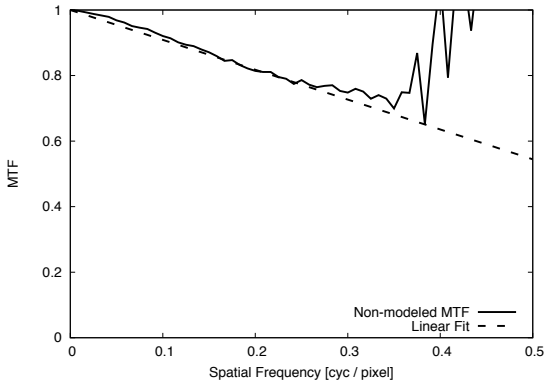
Dividing these two results will give the non-modeled MTF of the system

$$MTF_{\text{nonmod}} = \frac{MTF_{\text{ctrl,meas}}}{MTF_{\text{ctrl,mod}}} = MTF_{\text{coll}} \cdot MTF_{\text{lens}} \tag{5.6}$$

5.4.1 Results



(a) Measured and modeled MTFs for the control experiment.



(b) Non-modeled MTF with a linear fit

Figure 5.14: Results of the control experiment to determine the non-modeled MTF of the laboratory system

In order to obtain these results, ten images were taken through the circular aperture. The same region of interest was taken from each image and the MTF was computed independently. The MTF measurements were then averaged to obtain the measured MTF. The measured and modeled MTFs are overlaid in Figure 5.14. The spread in this plot was determined by the minimum and maximum values seen in the ten trials. Measurement noise became significant after 0.3 cycles per pixel. As such, the expression for the non-modeled MTF will only be fit on the data under 0.3 cycles per pixel. Since the sparse

aperture system's cutoff frequency is around 0.3 cycles per pixel, this noisy region of frequency space is irrelevant. Since there is little curvature in the data, and noise is present, a simple linear model was fit to the data, shown in Figure 5.14. The non-modeled MTF was approximated as

$$MTF_{\text{nonmod}}(\xi) = 1 - 0.911 \cdot \xi \tag{5.7}$$

where ξ is spatial frequency in units of cycles per pixel.

5.4.2 Simulated Experiment

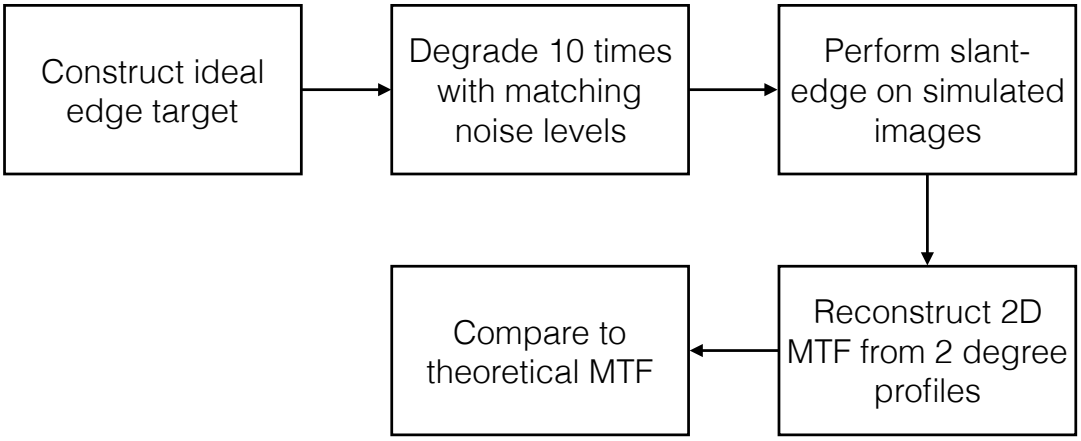
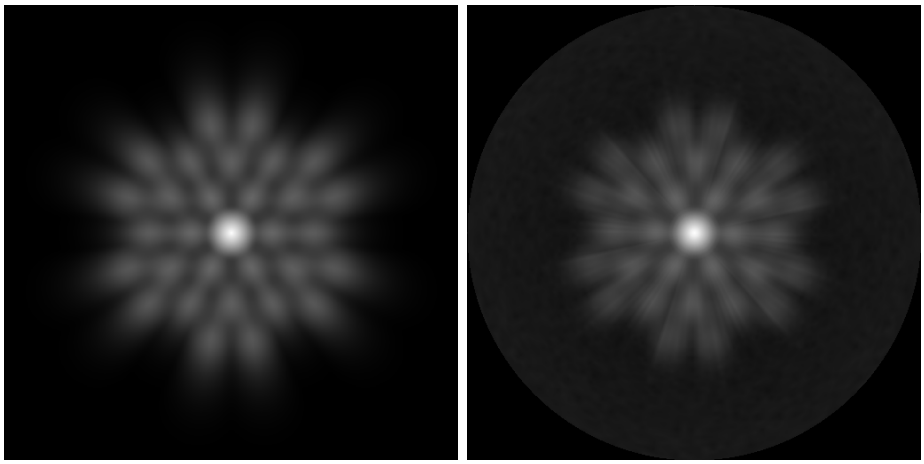


Figure 5.15: Algorithm for simulating the laboratory OTF validation experiment

Given the result in the previous section, the non-modeled MTF can be integrated into the model. Doing this would, in fact, make it a modeled MTF, however, it will continue to be termed the non-modeled MTF for consistency's sake. After integration, the entire laboratory experiment can be simulated in the sparse aperture model to see the expected results before performing the experiment in a laboratory. The modeling process is illustrated in Figure 5.15.

As can be seen in the figure, the modeling procedure closely mirrors the actual laboratory experiments, with the exception that image degradation is done digitally instead of optically. The effects of noise, non-modeled MTF and MTF measurements/reconstruction procedures should all be present in the simulated 2D MTF simulated measurement. The



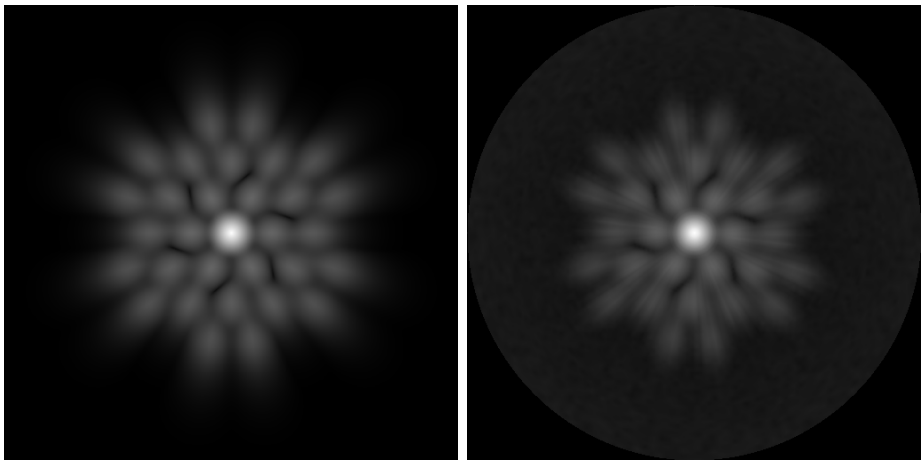
(a) Theoretical MTF (b) Simulated measured MTF

Figure 5.16: Results of simulated experiment for a Golay-6 aperture with no additional wavefront error.

results of the simulated experiment for an un-aberrated system are given in Figure 5.16. As can be seen, it is reasonable to expect to see all of the secondary peaks in this case. The tails on the outer peaks due to spectral averaging are unlikely to be detected due to the presence of noise, non-modeled MTF and the effects of the slanted-edge MTF technique. However, the shape, spacing and magnitude of the secondary peaks all appear to be reasonable goals for the laboratory experiment.

Figure 5.16 (b) has a noticeable disk around the simulated measured MTF. This is due to image noise and will appear on all MTF reconstruction derived from measured data. Image noise will propagate through the slanted-edge MTF measurement technique and manifest in noise in the reconstructed MTFs. While this noise is present at every frequency, it is most visible when the MTF is zero or close to zero, which occurs at the higher frequencies. This noise floor is at most 1% MTF is is only visible due to the contrast-stretching present in the figures.

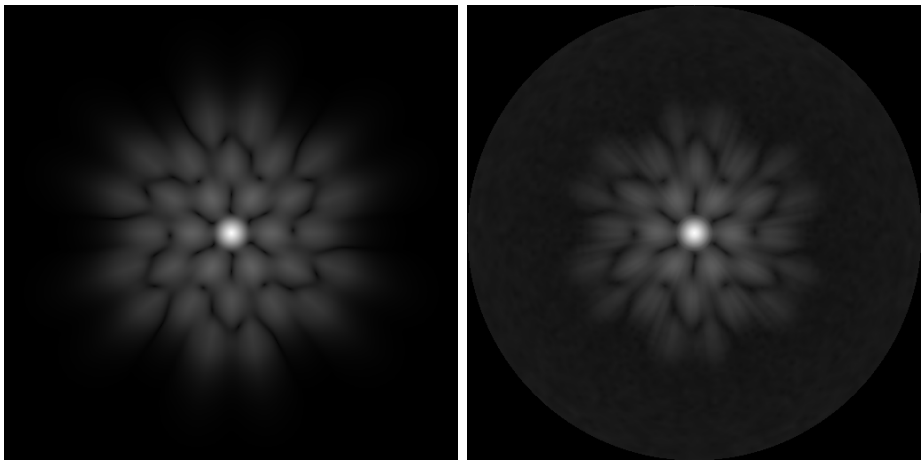
As wavefront error is introduced into the system, the MTF of the system degrades. This manifests in both decreased magnitude of secondary peaks, as well as exaggerated valleys between peaks. As was shown in Figure 5.12, at 0.5 waves of defocus, there are exaggerated



(a) Theoretical MTF (b) Simulated measured MTF

Figure 5.17: Results of simulated experiment for a Golay-6 aperture with 0.5 waves of defocus wavefront error.

valleys that spiral out from the central peak. In the case of 1.1 waves of defocus, there are exaggerated valleys between the peaks of the first ring, as well as a "star pattern" of valleys in the outer ring of peaks. The results of the simulated experiment for the case of 0.5 waves of defocus are shown in Figure 5.17. The spiral pattern can still clearly be seen; thus, it is reasonable to expect it to be observed in the actual experiment. In the case of 1.1 waves of defocus, shown in Figure 5.18, the valleys in the first ring of peaks are clearly visible. While the "star pattern" is visible the peaks outside of the pattern are starting to fade into the noise. If there is more noise present in the actual experiment (as it turns out there was), the "star-pattern" will be less visible in the measured data.



(a) Theoretical MTF (b) Simulated measured MTF

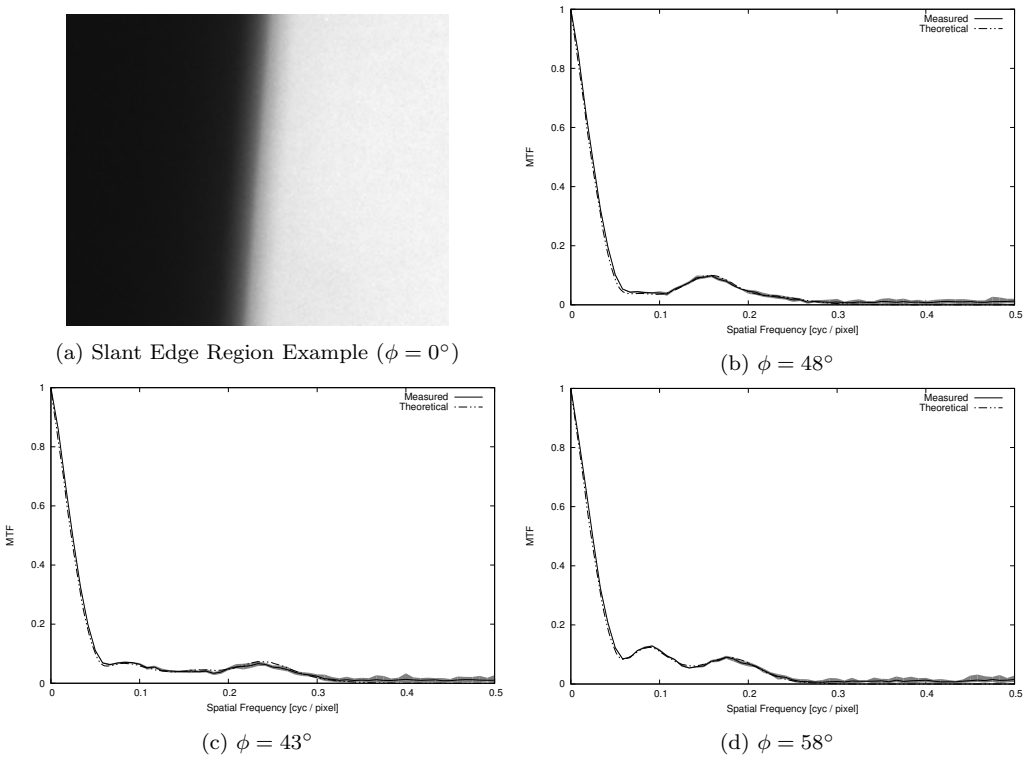
Figure 5.18: Results of simulated experiment for a Golay-6 aperture with 1.1 waves of defocus wavefront error.

5.5 Results

5.5.1 Un-aberrated System

The primary set of results for this study is the comparison of model MTF predictions to measured MTF as a function of the angle of the aperture. The model prediction for this experiment is shown in Figure 5.12 (a) and will be plotted with the measurements using a dashed line. In order to produce a model prediction, the $f/\#$ of the system is a key parameter. This is determined by the combination of the focal length of the lens and the encircled diameter of the Golay-6 pattern on the mask. In this experiment, the $f/\#$ of the system is 24. MTF predictions were calculated using Equation 5.2 over the bandpass shown in Figure 5.3 with no wavefront error. In addition, the nonmodeled MTF found in the control experiment is included into the model prediction.

In order to obtain the experimental results, the mask was rotated over a 60° region in 2° increments. Since the Golay-6 aperture has six-fold rotational symmetry, this is sufficient to capture all of the distinct radial profiles of the MTF. At each orientation, 10 images were acquired with a 0.25 second exposure. This exposure was determined experimentally to ensure the bright region of the image had digital counts approximately in the center of the detector's dynamic range, so that non-linearity issues would not be a concern. The



(a) Slant Edge Region Example ($\phi = 0°$)

(b) $\phi = 48°$

(c) $\phi = 43°$

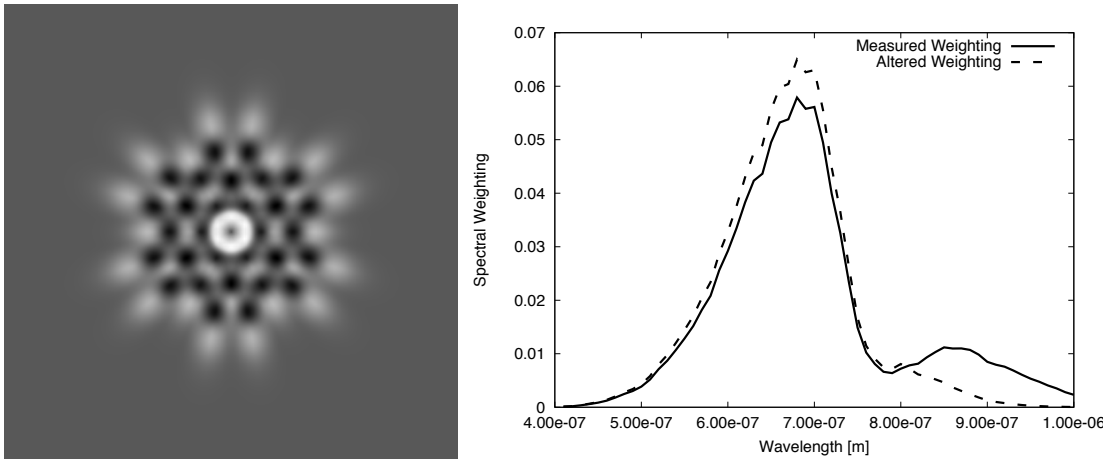
(d) $\phi = 58°$

Figure 5.19: Sample results for MTF validation study of an un-aberrated Golay-6 aperture

detector was set to cool to $-15°$C, although depending on room temperature the detector temperature ranged from $-7°$C to $-15°$C.

Due to the non-isotropic nature of sparse aperture OTFs, an alignment between model predictions and measured MTFs is necessary to perform any quantitative analyses. Fortunately, this is a simple task, due to the peaks that arise in the MTFs. By utilizing the angular markings on the rotation stage, it is known that measured MTF profiles are separated by approximately $2°$. However, due to the finite angular extent of the markings and manual error, error in the spacing of the samples was introduced. A two-step automatic procedure is used to perform the alignment. The first step holds the spacing at a constant $2°$ and varies the global offset between the model and measurements. This accounts for the orientation at which the mask was mounted on the rotation stage. Secondly, each profile is allowed to vary over a small region around the initial alignment to minimize

RMS error. This accounts for error in the angular position of each measurement. Since these measurements were performed in the dark, slight errors in precise alignment may have occurred.



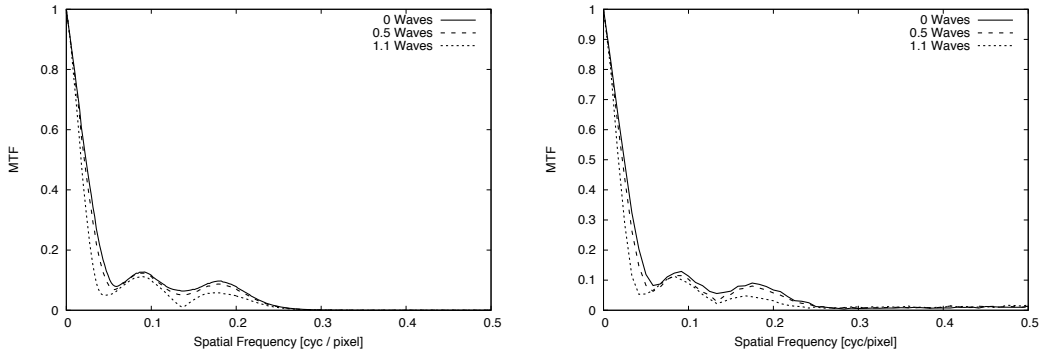
(a) Change in MTF due to a decrease in NIR transmission (RMS error is 0.29% MTF).

(b) Example of a potential error (decrease in NIR transmission) in the spectral weighting function.

Figure 5.20: Errors in the spectral weighting function change the positions and widths of peaks.

Results of the MTF validation study for an un-aberrated system are given in Figure 5.19. As can be seen, the measured and predicted MTFs match nearly exactly. RMS error will be given in units of MTF percentage, which is simply the MTF error multiplied by 100. The average RMS error between the measured and theoretical MTF was 1.07% over all angles, with a range of 0.67% to 1.35%. The error that does exist can be explained by a few factors. The first is imperfect knowledge of the spectral weighting function for calculation of MTF predictions. The spectral weighting function did not include the transmission function of the collimator optics, sparse aperture mask or lens optics. Since the spectral weighting function determines the radial blurring of the OTF, error in this function could explain any small discrepancy in the widths of the secondary peaks between the measured and theoretical MTFs. For example, Figure 5.20 shows an example where the NIR portion of the illumination was attenuated. In this case, peaks moved radially outward and decreased in width, leading to a change of 0.29% RMS difference.

While spectral weighting can account for error in the width and positioning of peaks, a portion of the error comes from image noise. Image noise propagates through the MTF measurement procedure and contributes to the error at every frequency. The measured MTFs in Figure 5.19 are the result of an average of 10 measurements and still have noticeable noise. The effects of image noise can clearly be seen in the region beyond 0.3 cycles per pixel, which is the cutoff frequency of the optics OTF. The measured MTFs are non-zero in this region due to image noise. Lastly, the precise $f/\#$ of the system is not known due to uncertainty in both focal length and diameter. A zoom lens is used in the system and while the 80mm focal length is nominally on one of the extremes, a geometric calibration would be required to determine the precise focal length. The effect of this error is similar to that of the spectral weighting function, in that it changes the radial position and extent of the MTF peaks.



(a) Model Predictions ($\phi = 0°$)

(b) Measured Data ($\phi = 0°$)

Figure 5.21: Sample results for MTF validation study of the Golay-6 aperture with various levels of defocus

5.5.2 Aberrated System

The modeled MTFs for a defocused system were shown in Figure 5.12 (b) and (c). The amount of defocus refers to the peak value of the wavefront error, which, in the case of defocus, occurs at the outer border of the aperture. As can be seen in Figure 5.21, the slope of the first peak in the MTF increases with worsening defocus error. In practice, this slope can be measured with high precision and exhibited no observable variance over the 10 measurements that were taken at each orientation. As such, this feature in the MTF

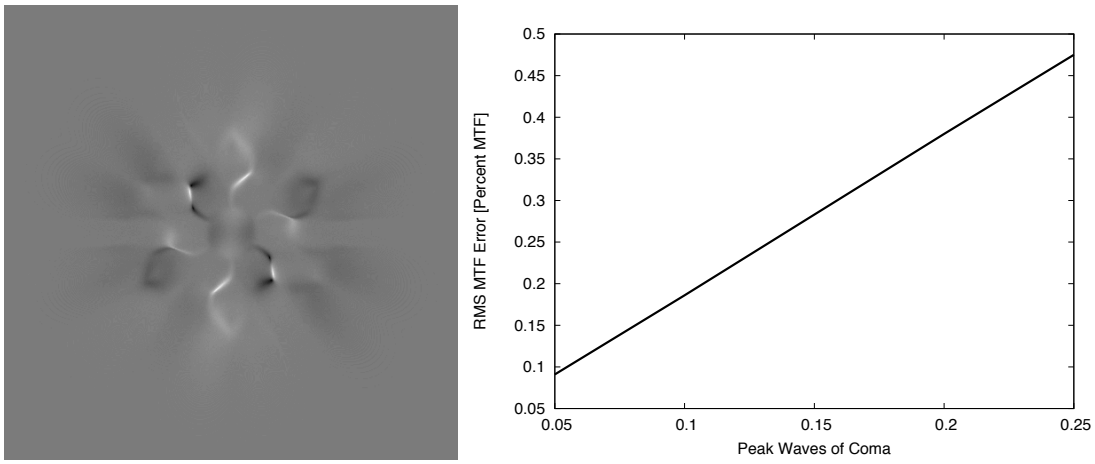
was used to determine the amount of defocus error present in the measurement.

Table 5.2: RMS errors between modeled and observed MTFs (units MTF%)

Angle (modulo 60°)	0 Waves Defocus	0.5 Waves Defocus	1.1 Waves Defocus
0°	0.86%	0.89%	0.88%
5°	0.67%	0.73%	0.55%
10°	0.97%	0.71%	0.58%
15°	0.92%	0.68%	0.55%
20°	0.99%	0.73%	0.65%
25°	0.97%	0.73%	0.61%
30°	1.13%	0.68%	0.71%
35°	1.22%	0.72%	0.63%
40°	1.22%	0.96%	0.48%
45°	1.17%	0.98%	0.69%
50°	1.25%	0.84%	0.71%
55°	1.35%	0.90%	0.82%

As can be seen in the figure, the results again match the behavior of the model very closely. For the data collected at 0.5 waves of defocus, the average RMS error across all angles was 0.79% MTF with a range of 0.68% to 0.98% MTF. For the data set collected at 1.1 waves of defocus, the average RMS error across all angles was 0.63%, with a range of 0.48% to 0.88% MTF. All sources of error that were discussed in the previous section also contribute to the error in this experiment, with image noise being a bigger issue on the 1.1 wave data set, due to the decreased magnitude of the secondary MTF peaks. The complete error results are given in Table 5.2. It is interesting that error decreases with increasing defocus. While this may be unintuitive, the overall magnitude of the MTF decreases with defocus, so, if error is truly only arising from the factors listed, such as error in the spectral weighting and $f/\#$, the magnitude of the error should decrease with the magnitude of the MTF. The only other observation that is apparent from the error data is that error was consistently high in the 35° to 45° range. This is to be expected, as those profile are in a range between peaks in the MTF, thus resulting in little MTF signal to be measured.

One final potential source of error is misalignment between the collimator and imaging



(a) Change in 0.5 wave defocus MTF from adding 0.1 waves of coma

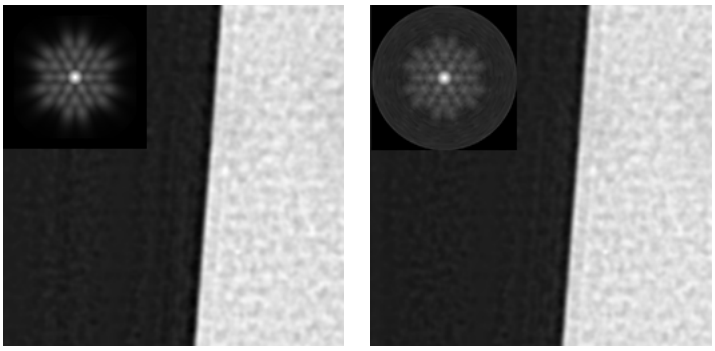
(b) RMS change in the sparse aperture MTF as a result of adding various levels of coma

Figure 5.22: Misalignment of the collimator and imaging system will result in coma, which can contribute to the error between the measured and modeled MTF

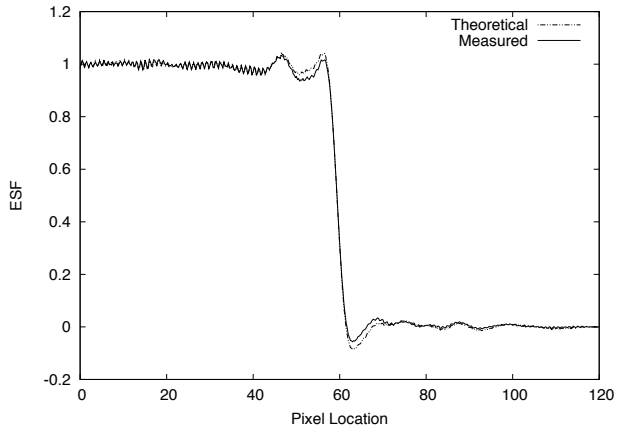
system, which can introduce coma into the measurement area. Confining analysis to the center of the image was supposed to mitigate coma from the imaging lens, however, coma from the collimator's parabolic primary mirror will still be present. Figure 5.22 shows the result of a simple sensitivity analysis of adding coma into the modeled MTF for the case of 0.5 waves of defocus. If an eighth of a wave of coma were introduced into the system, which is the quoted wavefront error of the collimator, then an RMS difference of 0.25% MTF would be introduced. This would explain a portion of the error seen in Table 5.2, along with spectral weighting error, non-modeled MTF and measurement noise.

5.5.3 Inverse Filtering and RER Comparisons

In order to put the experimental results into image quality terms, a brief study was done to determine the effect of the difference between model and measured MTFs on the relative edge response (RER). In the GIQE 4, the RER is measured after post-processing and thus, inverse filtering needs to be applied to the images. Given the images captured through the laboratory system, post-processing filters derived from both the measured MTF and predicted OTF were used to restore the image. RER was then measured on the restored images and the values were compared. An example result is shown in Figure 5.23.



(a) Restored with Modeled OTF (b) Restored with Measured OTF



(c) ESF Comparison

Figure 5.23: Results of the RER study for an unaberrated system ($\phi = 0°$, $\gamma = 0.01$)

In order to design a post-processing filter for the measured MTFs using Equation 5.1, a representation of the system's two-dimensional OTF is required ($H^*(\xi, \eta)$ and $|H(\xi, \eta)|^2$). However, in this situation, only one-dimensional radial profiles of the MTF, $|H(\xi, \eta)|$, are available. As such, bilinear interpolation was used in polar frequency space to fill out the two-dimensional MTF, $|H(\rho, \phi)|$, and was then converted back to ξ and η. Since no information about the phase transfer function can be measured using the slant edge method, the MTF was simply assumed to be the real part of the OTF and the imaginary part was set to zero.

As can be seen from Figure 5.23, the restored images and edge spread functions (ESFs) look similar, exhibiting the same ringing artifacts, with the modeled restoration filter exhibiting slightly more edge overshoot. For the unaberrated system, the average RER over all angles was 0.240 for the modeled OTF restoration and 0.232 for the measured MTF restoration. This equates to a ΔNIIRS of 0.04, well below the "just-noticeable" difference threshold of 0.1 [Fiete, 1999]. The average percent error between the two RER measurements was 3.33%, with the modeled always having the greater RER, although this was not a general trend and did not occur in the two aberrated data sets. As a tradeoff for the higher RER, there was also more artifacting for the modeled OTF restoration. This is in the form of a ringing artifact instead of a simple edge overshoot, so it does not have a method by which to include it in the ΔNIIRS calculation. For the 0.5 wave defocus data, the average modeled OTF restoration RER was 0.188 and the average measured MTF restoration RER was 0.185, a ΔNIIRS of 0.014. The average percent error over all angles was 1.20%. Finally, for the 1.1 wave defocus data, the average modeled OTF restoration RER was 0.169 and the average measured MTF restoration RER was 0.164, a ΔNIIRS of 0.04. The average percent error over all angles was 3.28%. Since the inverse filter relies on the OTF predictions, all error sources from the previous two experiments affect these results. In addition, the lack of phase transfer function knowledge and interpolation error will contribute to the error. Interpolation error is most noticeable toward the periphery of the measured MTF, where samples are less dense and there is noticeable noise present.

5.6 Conclusions

In this chapter, a validation study was presented for the sparse aperture image quality model set forward in Chapter 4. An optical design was presented that utilized a sparse aperture mask and an SLR lens to simulate a spare aperture telescope. The lens focus adjustment was then used to introduced controllable wavefront error into the system. A summary of the results is shown in Figure 5.24, which shows the observed results compared to the results of a simulated experiment using the sparse aperture model. Before looking at numerical errors, a number of conclusions can be drawn from a simple, visual comparison. Firstly, the distribution of secondary peaks matches between the model and measurement, indicating that the MTF calculation through autocorrelation is being correctly performed. Secondly, the shape and width of the secondary peaks matches, indicating that the spec-

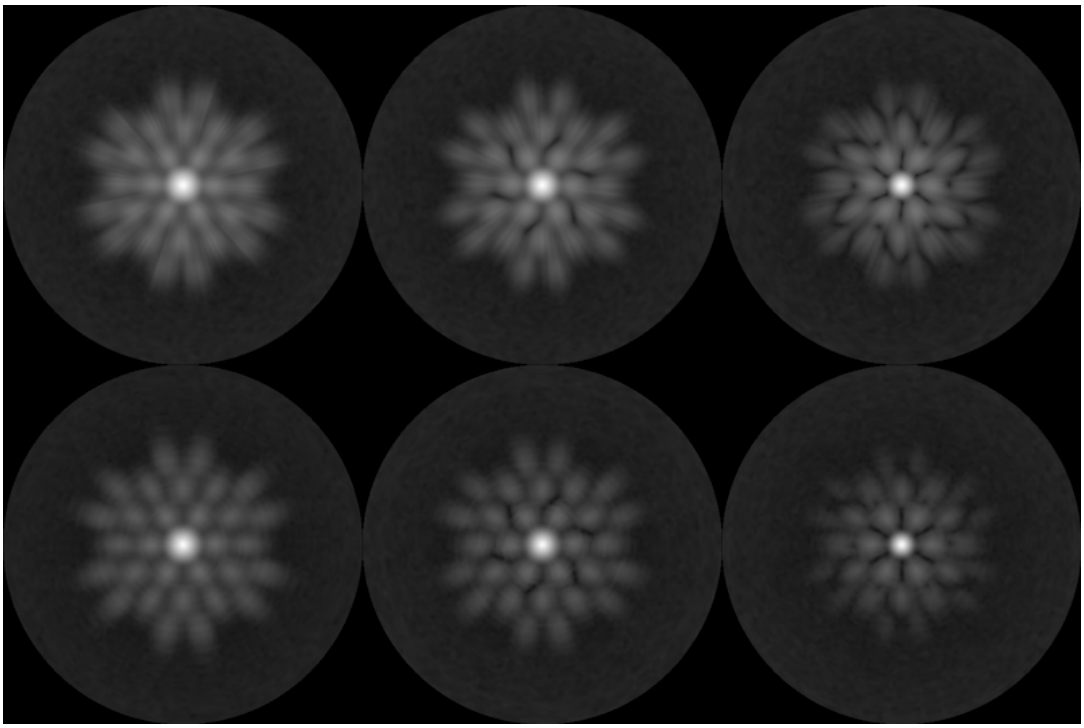


Figure 5.24: Two-dimensional MTF reconstructions from the simulated experiment (top) and the actual experiment (bottom) for 0 waves (left), 0.5 waves (center) and 1.1 waves (right) of defocus.

tral modeling is being performed correctly. Lastly, matching effects of wavefront error can be seen in the measurements. The spiral pattern in the 0.5 wave defocus data set and the central valleys in the 1.1 wave defocus data set are both clearly visible. The "star pattern" in the 1.1 wave defocus data set is visible, although the signal-to-noise ratio towards the outer peaks is lower in the measurement than in the simulated experiment. These matching patterns indicate that wavefront error is being properly accounted for in the OTF calculations. In Figure 5.24, it is also clear that the measured MTFs are slightly lower in magnitude than those in the simulated experiment, indicating that there was some additional non-modeled MTF that was not accounted for. In addition, the noise in the measured data had slightly more frequency structure than the noise in the simulated experiment. However, these sources of error are minor and do not affect the conclusions of the validation study.

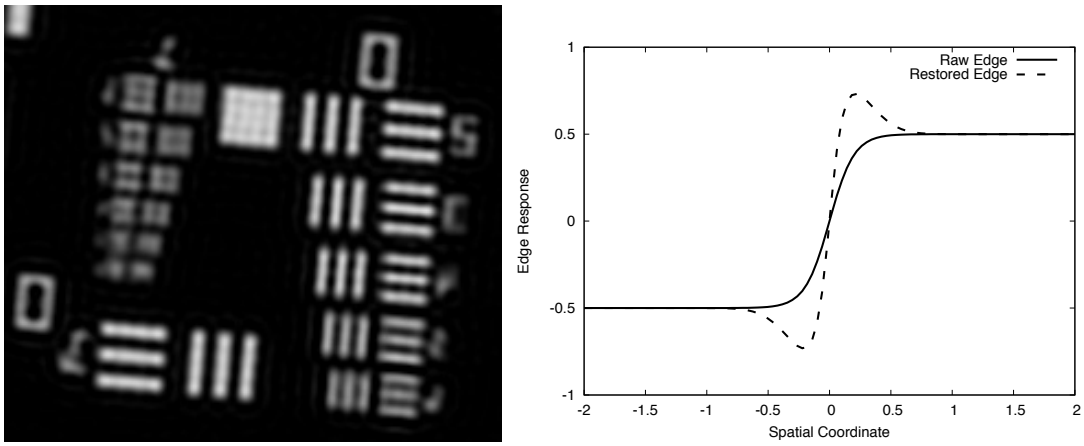
The visual comparison of the measured data with the model predictions gave a number of useful conclusions. To back these conclusions up numerically, a number of quantitative results were also produced. Without any wavefront error in the system, the RMS error between the measured MTF and the theoretical MTF was approximately 1% MTF. When 0.5 waves of defocus was introduced, the RMS error went down to 0.79% MTF and when 1.1 waves of defocus was introduced, the error went down to 0.63% MTF. Given that peaks in the MTF of the Golay are generally in the neighborhood of 10% MTF, these error values are relatively low. Additionally, post-processing filters derived from the modeled OTF and measured MTF were used to restore the measured images. The model's prediction of ringing artifacts were observed in real world data and the relative edge response measurements matched closely. At all levels of defocus, the ΔNIIRS between measured and modeled restored images due to RER was less than 0.04, where a 0.1 ΔNIIRS is considered a "just-noticeable" difference [Fiete, 1999].

While this validation study allowed for a number of conclusions on the validity of the model's OTF predictions, there are a number of limitations. The primary limitation is that only MTF was measured, so no conclusions about the phase transfer function (PTF) can be made. Also, due to the design of the laboratory system, the only controllable wavefront error was defocus. In practice, sparse aperture systems will have sub-aperture aberrations that are independent of one another. A laboratory design that facilitates this sort of wavefront error would be desirable. In addition, while ringing artifacts were observed and did match model predictions, they were relatively minor on such a simple edge target. As such, additional studies on more complicated targets were also performed and will be described later.

Chapter 6

Artifact Validation

6.1 Introduction



(a) A simulated restored image with a circular aperture.

(b) Effects of post-processing on edges in imagery taken with a circular aperture

Figure 6.1: Post-processing on traditional imagery makes imagery sharper and only results in simple edge overshoot artifacts.

Image artifacts, which arise as a result of post-processing, are a key factor to understanding image quality. For systems with monolithic aperture functions, these artifacts manifest as edge overshoots, as plotted in Figure 6.1 (b). However, under reasonable levels of restoration, these artifacts only have a small effect on image quality, as Figure 6.1 (a) shows. In general, imagery remains smooth and the effect is only noticeable close to edges.

Once the aperture shape becomes more complicated, these image artifacts become more complex and a larger factor in image quality. As was plotted in Figure 5.23 (c), when dealing with sparse aperture imagery, these artifacts can manifest as secondary peaks after the initial overshoot, resulting in false signals in the image. This has a larger negative effect on image quality than the simple edge overshoot artifact. If we consider the General Image Quality Equation,

$$\text{NIIRS} = 10.251 - a \cdot \log_{10}(GSD) + b \cdot \log_{10}(RER) + c \cdot H - d \cdot \frac{G}{SNR} \qquad (6.1)$$

there was a means by which to integrate artifacts into the image quality of traditional imagery [Leachtenauer et al., 1997]. This is done using the H term, which is the edge overshoot and the G term, which is the gain of the post-processing filter. For conventional imagery, it was found that these terms correlated linearly with the NIIRS rating, although the magnitude of their coefficients was much less than the terms on GSD and RER. However, for systems with non-conventional apertures, it is unclear how to quantify these artifacts, what correlation they have with image quality and how important they are relative to other image quality terms.

While the top-level goal of characterizing the effects of generic artifacts on image quality is beyond the scope of this research, there are some prerequisite challenges that need to be solved before the image quality challenges can be addressed. Since constructing and launching a sparse aperture telescope would currently be an extremely challenging and expensive venture, gaining an understanding of image quality tradeoffs with simulated data would be desirable. Much of the previous work mentioned in Chapter 4, particularly [Block, 2005], has predicted the existence of artifacts in restored imagery. This research will take the model predictions of artifacts and attempt to reproduce them in real data collected with the sparse aperture system described in Chapter 5. However, it will be left to future investigators to address the impact of these artifacts on image quality.

6.1.1 Causes of Artifacts

Since artifacts only arise in imagery that has been restored, they must occur due to imperfections in the restoration process. As has been described in Chapter 4, image

restoration is performed with a Wiener filter of the form,

$$\hat{F}(\xi,\eta) = W(\xi,\eta)\cdot G(\xi,\eta) = \frac{H^*(\xi,\eta)}{|H(\xi,\eta)|^2 + \gamma\cdot|S(\xi,\eta)|^2}\cdot G(\xi,\eta) \tag{6.2}$$

In this equation, $G(\xi,\eta)$ is the measured raw image spectrum (with noise), γ is a user-set smoothness parameter and $S(\xi,\eta)$ is a deterministic spectrum of a smoothness operator. While noise in the image limits the effectiveness of restoration, the only term in which error can arise is $H(\xi,\eta)$, which is the OTF of the imaging system. The OTF of the system varies spectrally, however, due to the spectral integration of the system, restoration cannot be performed at each individual wavelength and must instead be performed on the integrated signal. In order to account for this limitation, the "effective" OTF of the system is used, which is computed as

$$H_{\text{eff}}(\xi,\eta) = \frac{\int_0^\infty SW(\lambda)\cdot H(\xi,\eta,\lambda)\ d\lambda}{\int_0^\infty SW(\lambda)\ d\lambda} \tag{6.3}$$

This equation gives insight into one potential cause of error, which will result in an increase in artifacts. $SW(\lambda)$ is the spectral weighting function of imaging system. This is the product of the spectrum of the incoming radiance, which spatially varies, the transmission spectrum of the optics and the responsivity of the detector in a given band. Error or lack of knowledge in any of these spectra will result in error in the spectral weighting function. Performing image simulations can show that errors in the spectral weighting functions should result in an increase in artifacts. This is shown in Figure 6.2. As can be seen in the figure, not accounting for the illumination spectrum of the scene will still allow for sharp edges in the resulting images, but at the cost of some contrast and artifacts. Using a gray-world assumption over the bandpass will result in an ineffective restoration, as the resulting edges will not be sharp and will exhibit large amounts of overshoot. For realistic imagery, the illumination spectrum will vary over the field of view. This leaves two options for realistic systems: accept these artifacts or somehow obtain an approximation of the incoming spectral radiance, such as with a coaxial low-resolution hyperspectral system.

However, the spectral weighting function is not the only term that can introduce error into the effective OTF. Error can also arise in the estimate of the monochromatic OTF,

(a) Image restored with true spectral weighting function

(b) Image restored with detector QE spectrum, but not illumination spectrum

(c) Image restored with a gray-world assumption. This image had large overshoots and required contrast mapping.

Figure 6.2: Artifacts can arise from incorrect or incomplete knowledge of the system's spectral weighting function.

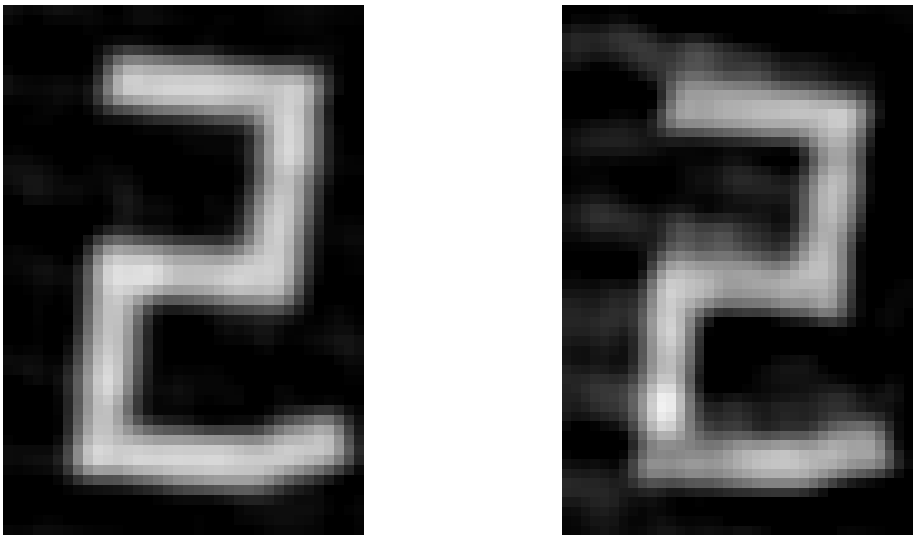
which was computed as the autocorrelation of the pupil function,

$$H\left[\xi,\eta;z_1=\left(\frac{1}{f}-\frac{1}{z_0}\right)^{-1}\right]=\frac{1}{k}\cdot p[-\lambda z_1\xi,-\lambda z_1\eta]\bigstar p[-\lambda z_1\xi,-\lambda z_1\eta] \tag{6.4}$$

Thus, any error must arise from error in the estimate of the pupil function, which is given as,

$$p[x,y,x',y']=|p[x,y,x',y']|\cdot e^{2\pi iW(x,y,x',y')} \tag{6.5}$$

Any error in the pupil function must come either from error in the magnitude or phase of Equation 6.5. Error in the magnitude would correspond to construction error in the telescope, such as mispositioned subapertures. Error in the phase would correspond to imperfect wavefront error estimates, which is likely to be the more common error. Just as with the spectral weighting function, image simulations will also verify this intuition. Figure 6.3 shows a simple simulation result. The image simulation was performed under -0.4 peak waves of coma and restoration was performed with and without this coma incorporated into the restoration filter. As can clearly be seen, wavefront error that has not been compensated for in the restoration filter should be another key source of



(a) Image restored with knowledge of the -0.4 waves of coma present in the raw image

(b) Image restored without any knowledge of wavefront error

Figure 6.3: Artifacts can arise from incorrect or incomplete knowledge of the system's wavefront error.

artifacting. Additionally, for aberrations such as coma, the restoration filter must be varied over the image plane to compensate for the spatially-changing aberration. The effects of both wavefront error and spectral weighting will be examined in this chapter.

6.2 Laboratory System

As in Chapter 5, a small-scale sparse aperture system in the laboratory will be used to match artifact predictions. In the slanted-edge experiments, post-processing artifacts were observed in the data, however, their magnitude was small and it would be desirable to see artifact manifestation in more realistic imaging scenarios. As such, a new target was needed for the laboratory collimator. A United States Air Force (USAF) 1951 tri-bar target was selected for this purpose. This is a traditional resolution target that is used in the remote sensing community. It was chosen due to the repeated spatial patterns, as well as the ability to automatically detect and analyze the imagery with image processing techniques.

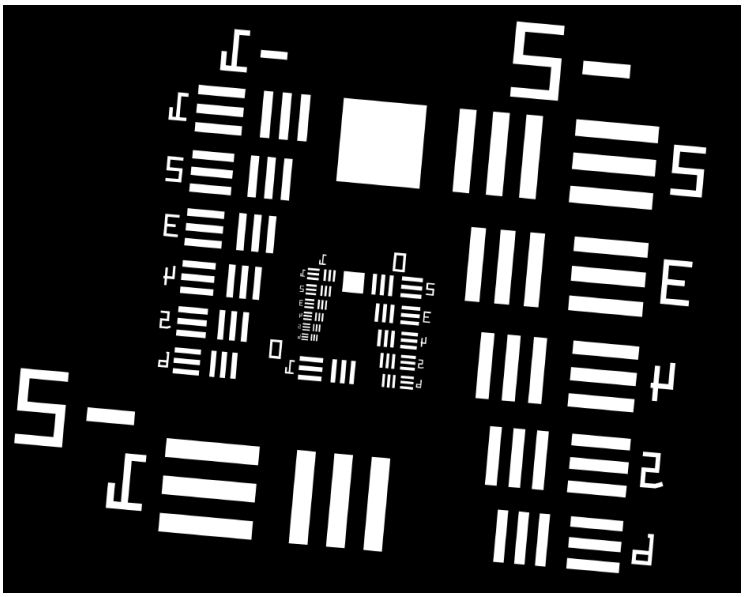


Figure 6.4: Digital representation of the USAF-1951 tri-bar target used in artifact validation experiments.

The target was fabricated by Applied Image Inc. as a chrome-on-glass print. A grayscale representation of the target is given in Figure 6.4. The tri-bars are transmissive regions, so the detected images with have a dark background with a bright foreground. The target was illuminated by the same integrating sphere as was used in Chapter 5. The illumination spectrum was plotted in Figure 5.6.

6.3 Methodology

The goal of this study is to validate the artifact predictions of the model. As such, procedures need to be designed to compare artifacts in measured and modeled data. This involves creating the same imaging scenario in the model and in the laboratory, as well as matching restoration procedures. Once these conditions have been met, qualitative observations and comparisons may be made about the data. Then, measurements must be performed on the data to extract quantitative information about the artifacts. Finally, some quantitative comparisons must be made between the measured and modeled data.

6.3.1 Model Configuration

In Section 6.1.1, it was shown that artifacts arise as a result of errors in the restoration filter. Thus, in order to observe artifacts in larger magnitudes than those on the slanted-edge data, the restoration filter must be slightly mismatched from the true imaging scenario. This is actually unavoidable, as there is wavefront error inherent in the laboratory setup. These errors include coma from the parabolic primary mirror of the collimator, defocus from the target potentially not being precisely located at the collimation plane, as well as the non-modeled MTF of the back-end imaging system. Unfortunately, it was not possible to characterize the wavefront error of the collimator due to lack of the necessary equipment. Seeing as the collimator and camera system were not precisely aligned, it would also be impractical to compensate for coma, as it would be difficult to determine the location in the image of the optical axis of the collimator.

Given the sources of wavefront error just described, an imaging scenario was designed to produce the modeled imagery. Both defocus and coma were added to the Golay-6 MTF. These aberrations were centered on the center of the image, with the acknowledgement that this is likely not exactly true due to misalignment of the imaging system and the collimator. The level of these aberrations was adjusted manually to roughly match the levels of artifacting seen in the measured images. Restoration was performed without any wavefront error compensation, giving rise to the necessary mismatch in the restoration filter. To further exacerbate the artifacts for the wavefront error study, the lens was slightly defocused. The lens was placed at optimal focus for the spectral weighting study.

6.3.2 Profile Extraction

Once images have been modeled or measured and restored, measurements must be extracted from the images. In this study, one-dimensional profiles over each tri-bar group will be extracted. Since the pattern was slightly inclined, as shown in Figure 6.4, the edge spread function reconstruction procedure from the slanted-edge MTF algorithm [Burns, 2000] can be used to extract profiles from the image over the tri-bar groups. If desired, these profiles can be extracted at a higher resolution than the image's native resolution. The implementation of this procedure written for this research automatically selects this resolution based on the orientation of the target and size of the input data. The majority of profiles in this study were extracted at a resolution of 4 samples per pixel.

In order to correctly position the profile measurements, the USAF-1951 target must be recognized and fit to a model. The pattern exhibits a number of detectable features. However, artifacts in the image will cause issues in detecting the bars, so a custom algorithm needed to be developed, which will be described here. Typically, the first step in a USAF-1951 recognition algorithm would be to binarize the image. However, using a single threshold on restored sparse aperture imagery would result in the artifacts also being classified as part of the foreground. Without a relatively clean segmentation of background and foreground, subsequent recognition steps would not be able to function. As such, an adaptive threshold was used to binarize the image. The neighborhood size was set to twice the width of the largest bar in the pattern. The output of this operation is shown in Figure 6.5.

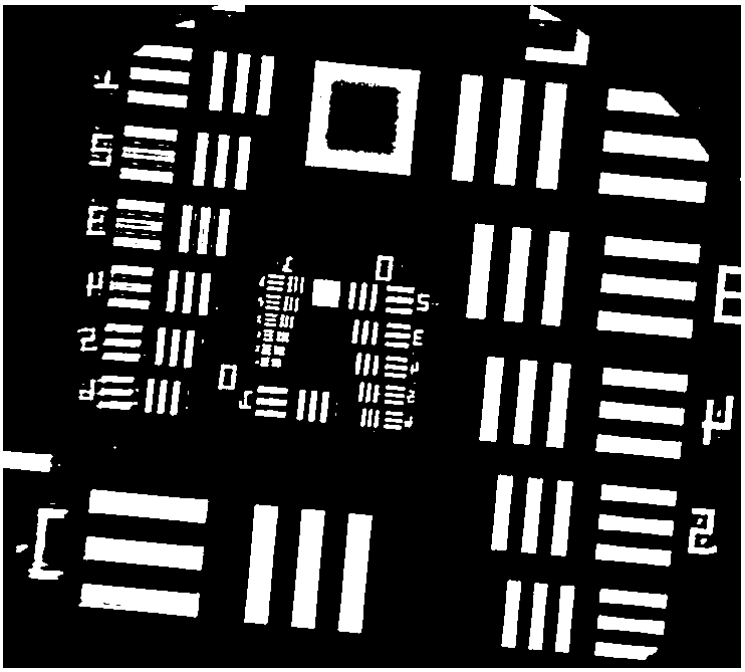


Figure 6.5: An adaptive threshold was necessary to isolate the bars from the post-processing artifacts.

As can be seen in the figure, the adaptive threshold can generally isolate the bars

from the artifacts. Once isolated, a connected components analysis was performed on the binarized image. This produces a connected component label image, as well as statistics, most importantly the area, for each connected component. The next task is to determine which of these connected components are the bars that compose a tri-bar group. This was done by looking at the aspect ratio of the connected component. Since the three bars in a group are equally spaced and the group itself is contained in a square, the aspect ratio of each bar is 5:1. A principal components analysis was performed on the pixel locations of each connected component analysis. The ratio of the square roots of the eigenvalues was used as an approximation of the aspect ratio. If the aspect ratio was close to 5:1, then the connected component was accepted as a bar. The results of this procedure are shown in Figure 6.6. This procedure was generally successful, although it could run into issues if artifacts were not successfully filtered out by the adaptive threshold, which did occur on images with higher levels of artifacting.

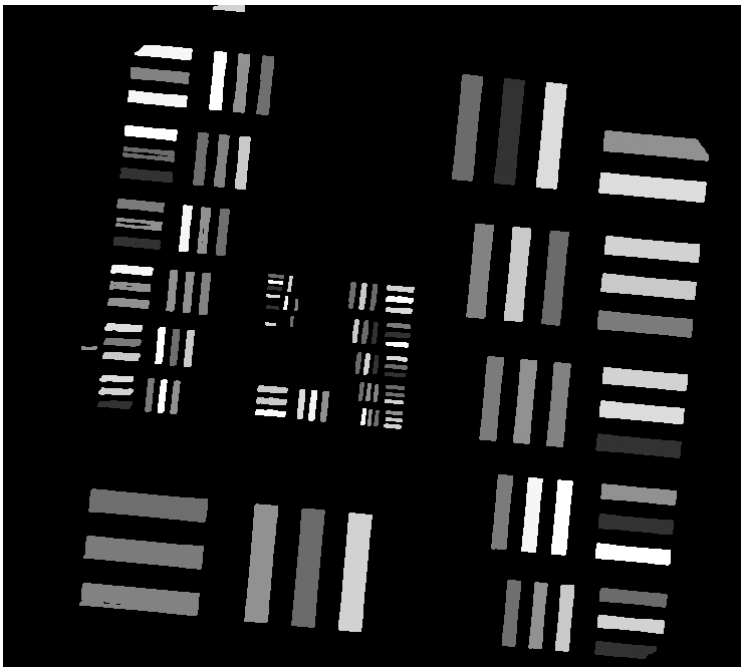


Figure 6.6: Analyzing the aspect ratio of the connected components was a useful step in filtering for the bars.

Once bars in the image have been identified, they can be used to identify the orientation of the target. By the design of the target, it is known that only two orientations of bars may exist in the image and those exist orthogonal to each other. As such, the orientation of each bar, which was already computed through the principal components analysis, was used to cluster the bars into oriented groups. A threshold of 10 degrees was used to detect outliers, which is useful to eliminate some false alarms from the bar detection phase of the algorithm. This may seem to be a loose threshold, however, orientation of the smaller bars may be noisy and inverse filtering will slightly alter the orientation of some of the smaller bars. Once divided, each orientation group can then be used to compute an average orientation vector, thus providing a set of two basis vectors for the target. Due to the issues with orientations of the smaller bars, bars were sorted by area and only the larger half were used in the basis vector computation.

Given the orientation of the target, the tri-bar groups can then be detected. The key insight here is that a tri-bar group is a group of three parallel bars of similar area that are equally spaced and have a spacing of one bar in between them. The centers of the three bars are also collinear. With this large set of constraints, a simple greedy algorithm was able to successfully group the tri-bars, along with eliminating the remainder of the false alarms from the bar detection phase. With the tri-bar groups detected, bounding boxes could be fit around the tri-bar groups that were aligned to the basis vectors of the target.

As was seen in Figure 6.6, not every tri-bar group will be able to be detected. While it is not necessary to detect every group, it is critical to know which groups were missed. From the design of the target, the spatial frequency of each tri-bar is

$$f = 2^{\mathrm{G}+\frac{\mathrm{E}}{6}} \text{ [lp/mm]} \tag{6.6}$$

where G is the group number and E is the zero-based index of the tri-bar in the group. Looking at two adjacent tri-bars in the same group

$$\frac{f_{i+1}}{f_i} = \frac{2^{\mathrm{G}+\frac{\mathrm{E}+1}{6}}}{2^{\mathrm{G}+\frac{\mathrm{E}}{6}}} = 2^{\frac{1}{6}} \approx 1.12 \tag{6.7}$$

That is, between every tri-bar group, the frequency increases by a factor of roughly 1.12. In spatial units, this means that the spacing between bars, or the side length of the bound-

ing box, decreases by a factor of roughly 0.89 between each group. Alternatively, the area of the bounding box should decrease by a factor of roughly 0.79 between each tri-bar group. Once the bounding boxes are detected, they are sorted in decreasing order of area and ratios between consecutive areas are computed. If the ratio is closer to the square of the expected ratio of 0.79 than the ratio itself, it can be inferred that a tri-bar group was missed and a blank is inserted until all misses have been detected.

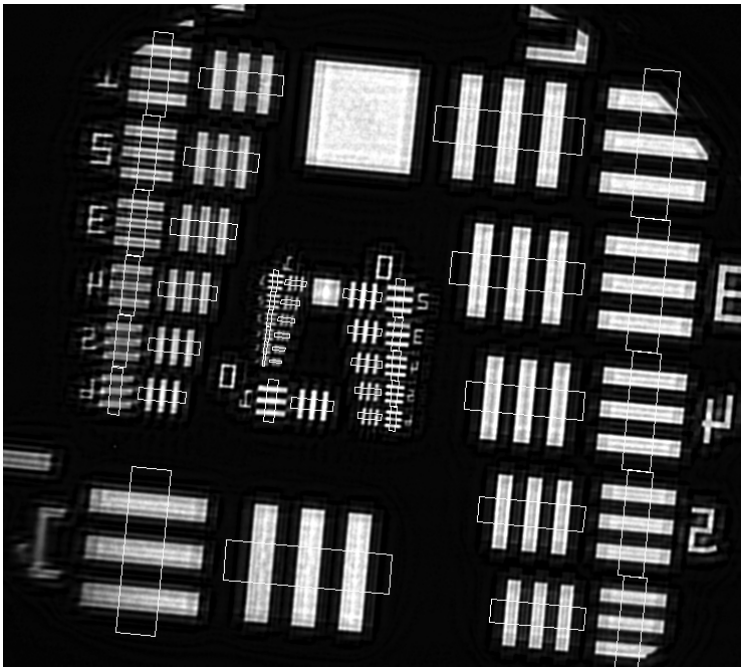


Figure 6.7: One dimensional profile regions are placed over each tri-bar target after recognition.

Once missed tri-bars have been detected, each detected bounding box can be associated with a group and element index. If in a given element, only one of the horizontal and vertical groups was found, the other can be inferred. Much like bounding box area, the spacing between the horizontal and vertical tri-bars falls off as a geometric series. So, the offset between the found tri-bar and the missing one can be inferred from the geometric series. It is also known that the horizontal tri-bars are always on the outside of the target. This can be used to determine the sign to attach to the horizontal basis vector for the

offset direction.

After completing any incomplete tri-bar pairs, the last step of the algorithm is to infer the location of any remaining missed tri-bars. By design, the USAF-1951 has nested copies of itself. That is, the inner groups of tri-bars are scaled replicas of the outer groups. Thus, the approach taken here will be to find a scaling and translation transformation between the outer and inner replicas of the target. The RANdom SAmple Consensus (RANSAC) model-fitting algorithm [Fischler and Bolles, 1981] will be used for this purpose. Once a transform between the two replicas of the target was found, the detected bounding boxes were used to infer the locations of the missed bounding boxes. Finally, a one-dimensional profile region was calculated from the bounding box by selecting the center portion and adding some padding beyond the ends of the box. The profile regions can be seen in Figure 6.7. As was mentioned earlier the projection technique from the slanted-edge MTF measurement algorithm is used to extract profiles at a resolution of 4 samples per pixel.

6.3.3 Error Metrics

Once profiles have been extracted from the restored USAF-1951 measured and modeled imagery, some error metrics must be extracted from the profiles in order to characterize how well the artifacts are being modeled. In order to calculate such metrics, the profiles will need to be co-registered. This was done manually, as efforts to automatically perform registration did not yield results with ample precision. Given two registered profiles, there are a number of metrics that may be defined. In this study, there are two large questions that need to be answered. The first is whether or not the model is predicting exactly the correct artifact pattern. The second is whether the model predicts artifacts that have equivalent effect on image quality.

It is unlikely that the model will produce the exact artifacts, as the actual imaging conditions cannot be exactly reproduced in the model. This is due to a number of factors. Firstly and most importantly, there are a number of uncharacterized sources of wavefront error in the system, namely defocus and coma from the collimator. The decentration between the collimator and imaging system was also unable to be characterized. Pixel phasing can also be a significant issue on the smaller tri-bar targets. The model requires a bitmap image as input and as such, assumes a certain pixel phasing, which may or may not

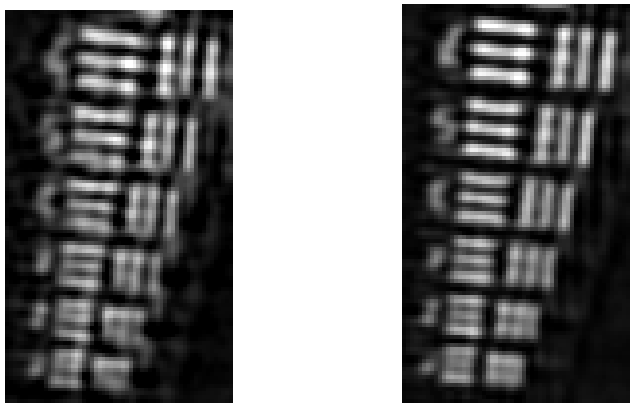


Figure 6.8: Pixel phasing can have a large effect on high-frequency targets. Notice the difference in modulation and bar spacing between the two pixel phasing instances shown here. Artifacting differences are due to other factors.

match the pixel phasing seen in the real imagery. This will affect the model's accuracy on all tri-bars, but will be most notable on the smallest groups, as it will affect the observed modulation. This effect is shown in Figure 6.8. Finally, the phase transfer function, which is the phase of the OTF, was not characterized for the laboratory imaging system, which again may cause some differences on the higher frequency tri-bars. In order to quantify the match between measured and modeled data, an average L1 distance between profiles will be used. RMS error is a more traditional choice, however, it was very sensitive to slight misregistration errors in the profiles.

The more interesting of the two questions is whether the model produces artifacts that have an equivalent effect on image quality. Given the sources of error, this is a more realistic goal. However, given that there is no equivalent to GIQE for sparse aperture imagery, there is no definitive metric for this question. For traditional imagery, the shape of the edge-overshoot artifact, which was plotted in Figure 6.1 (b), is generally constant. As such, only the peak magnitude of the overshoot was necessary. However, in sparse aperture imagery, the shape is not constant and depends on subaperture layout, orientation and spatial frequency, among other factors. As such, the peak magnitude alone is not sufficient. In this research, the area of the artifact will also be measured. Area alone is not ideal either however, as it also discards shape information. Thus, peak and area will both be measured and compared. An illustration of these two measurements is given in Figure 6.9. These

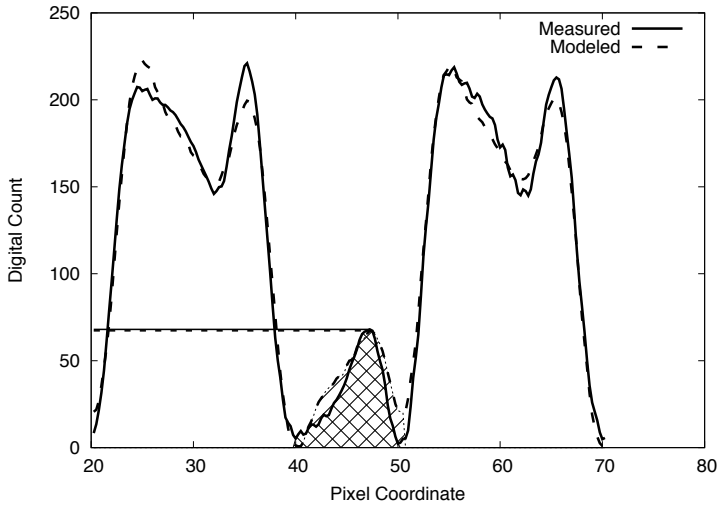


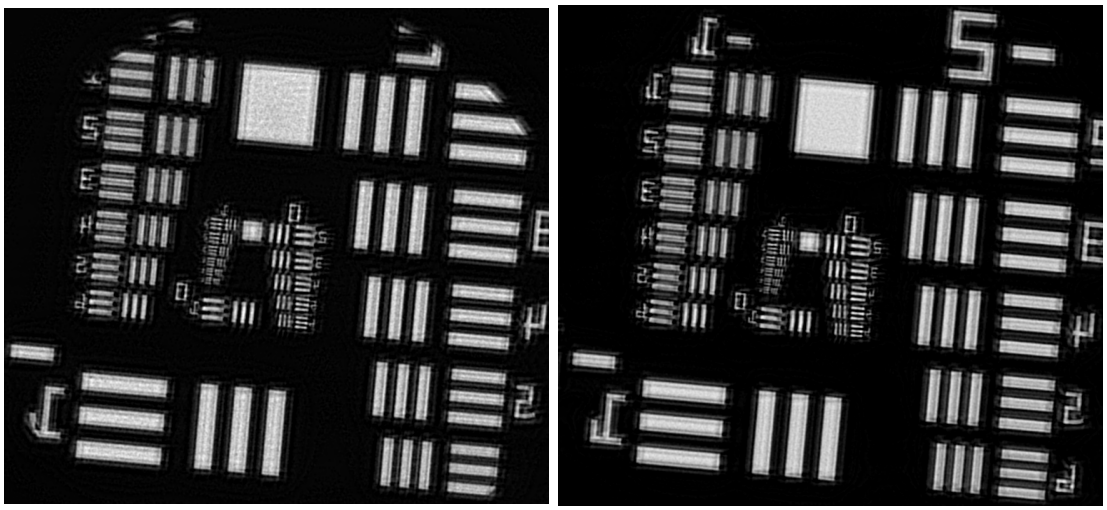
Figure 6.9: Illustration of artifact area and peak measurements

measurements will be used in this study with the acknowledgement that further research is needed to correlate them with image quality and that other measurements might be more predictive of image quality.

6.4 Wavefront Error Results

As was described earlier, the laboratory sparse aperture system was used to capture an image of the USAF-1951 target. This was done under the inherent wavefront error of the setup, along with a slight defocus of the imaging lens to increase artifacting. The imagery was then restored without any wavefront error compensation. A zoomed-out view of the measured and modeled restored imagery is shown in Figure 6.10.

From Figure 6.10, it is difficult to tell how well the artifacts match, however, one of the key differences that will explain the quantitative results is visible. While the noise levels in the two images is actually identical in terms of standard deviation, the measured data appears to the human eye to be noisier. This is due to the presence of spatial correlation of noise in the measured data. The noise power spectra for both images are shown in Figure 6.11. It is clear from the figure that the measured noise spectrum exhibits different behavior than that of the modeled data. The low-frequency peak in the



(a) Measured restored image

(b) Modeled restored image

Figure 6.10: Imagery results of the wavefront artifact validation experiment.

measured spectrum is slightly wider and the noise power is relatively constant over the range of $\pm 0.05 - 0.15$ cycles per pixel, whereas the modeled spectrum falls off in this range.

This discrepancy in the behavior in low-frequency noise is easily explainable. In the model, noise is simulated as the sum of photon and detector noise. However, each of these effects is modeled independently at each pixel, leading to spatially-independent noise and relatively lower levels of low-frequency noise. Since the human eye is more sensitive to lower frequencies, the spatially-correlated noise in the real data will have a larger perceived image quality impact than the spatially-uncorrelated noise. In addition, this structured noise also affects the artifacts, which can contribute to some the discrepancies in the quantitative metrics. Adding support for structured noise is a future improvement that can be made to the model. This would require a more sophisticated characterization of the noise performance of the detector being modeled.

For a closer look at artifacting, Figure 6.12 shows an enlarged version of bars 8-11, which are the last 4 bars in the image's top-left group. Bar groups are numbered in decreasing order of size, starting from 0. Being towards the edge of the field of view, the effects of coma can clearly be seen in this figure, through the matching directional smear

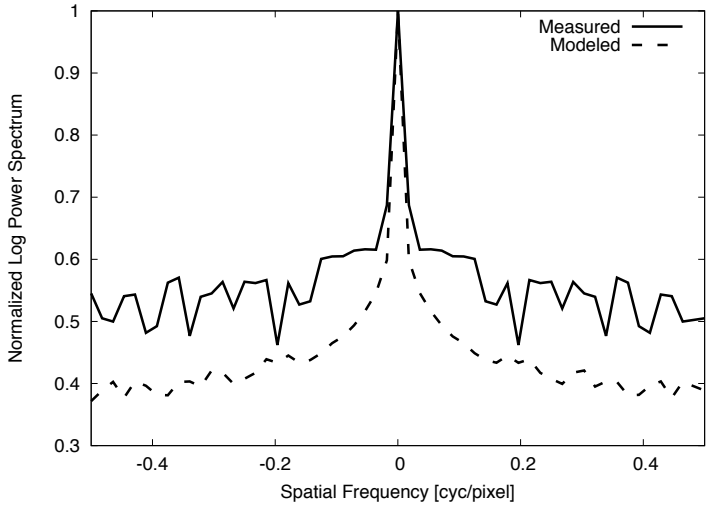
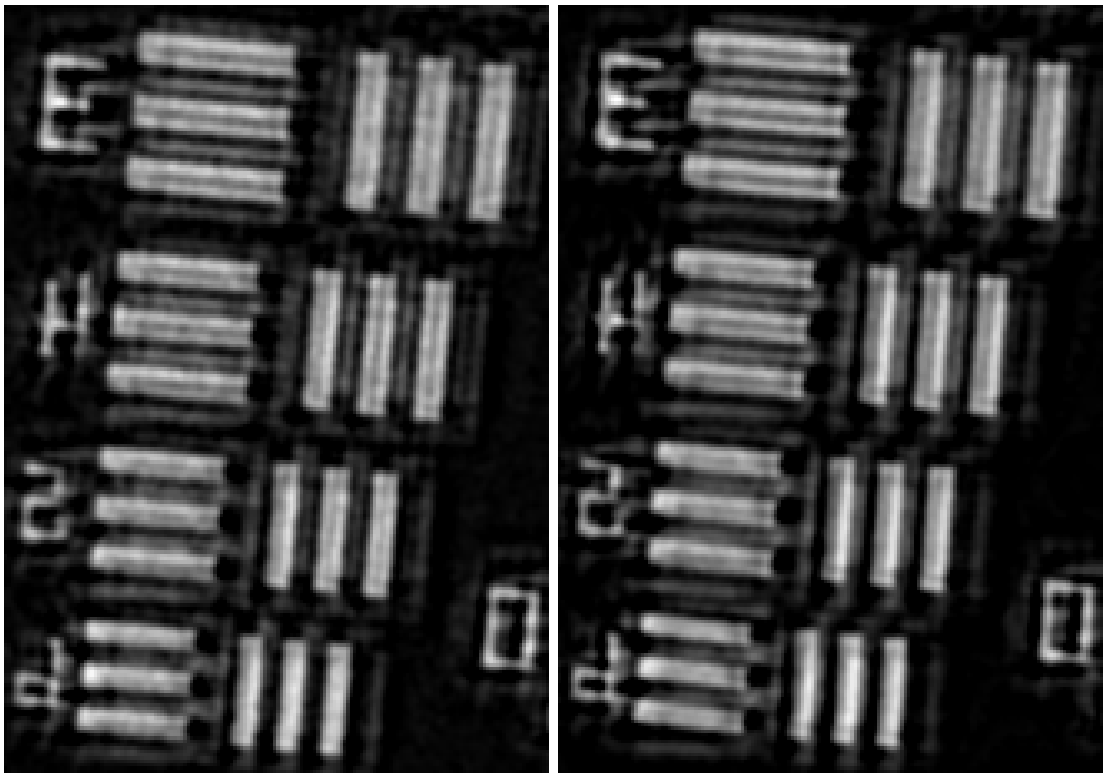


Figure 6.11: Relative log noise power spectra in the measured and modeled raw imagery. The RMS of the noise is equivalent in both images.

on the numbers next to the tri-bars. While not explicitly shown here, flipping the sign on the coma aberration will shift these artifacts to the other side of the numbers. Artifacts appearing between horizontal bars are also positioned correctly, as they are slightly shifted to the right, relative to the bars in both measured and modeled imagery. Vertical bar artifacting is less visible in both images, as artifacts are located closer to the bars, making them look like a simple blur. The smallest vertical group is interesting, as the artifacts merge with the bars in both images, resulting in a bright region inside the three bars. The artifact under the last bar in each horizontal group also shows matching behavior, as it is consistently not parallel to the bar in either image. This figure also shows the effect of structured noise, as the artifacts in the modeled imagery tend to be slightly smoother than those in the measured imagery.

Figure 6.13 shows an enlarged version of the inner level of the tri-bar target. These tri-bars more clearly show off the non-parallel artifact, which can be seen above the horizontal tri-bars on the right side of the images. The highest frequency tri-bar group also shows some interesting matching behavior. Looking at the vertical tri-bars in group 18-21 shows that the spacing between the center and left bars is smaller than that between the center and right bars. This effect also occurs in the horizontal tri-bar in group 22 and 23 in

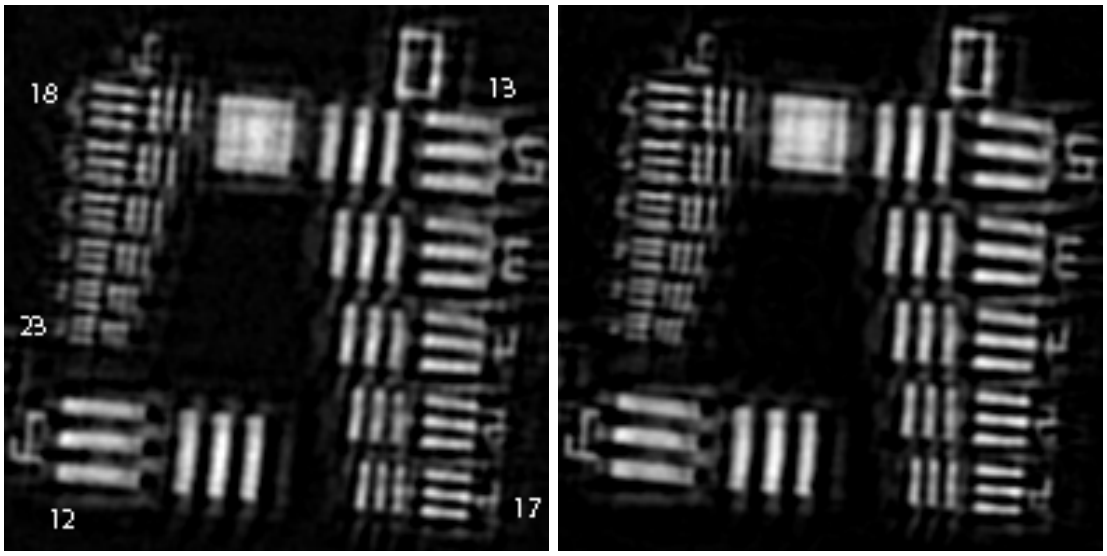


(a) Measured restored image (b) Modeled restored image

Figure 6.12: Enlarged version of bars 8-11 in the wavefront artifact validation experiment.

both measured and modeled images. This occurs due to matching pixel phasing between the input bitmap to the model and the actual imaging scenario. However, this occurred by chance and as shall be seen in the spectral results, is by no means guaranteed. Since this level of the target is closer to the center of the image, the effects of coma should be less pronounced. This is indeed the case, as the numbers on the right side of the images do not exhibit the same streaking that was seen in Figure 6.12.

Figure 6.14 shows two sample registered profiles from the tri-bars imagery. These profiles provide another layer of confirmation that the model is predicting artifacts that show up in the actual imagery. The measured and modeled curves have very similar shapes and exhibit peaks in roughly the same locations. While the precise positioning and magnitude varies slightly, the modeled and measured artifacts are definitely comparable.



(a) Measured restored image (b) Modeled restored image

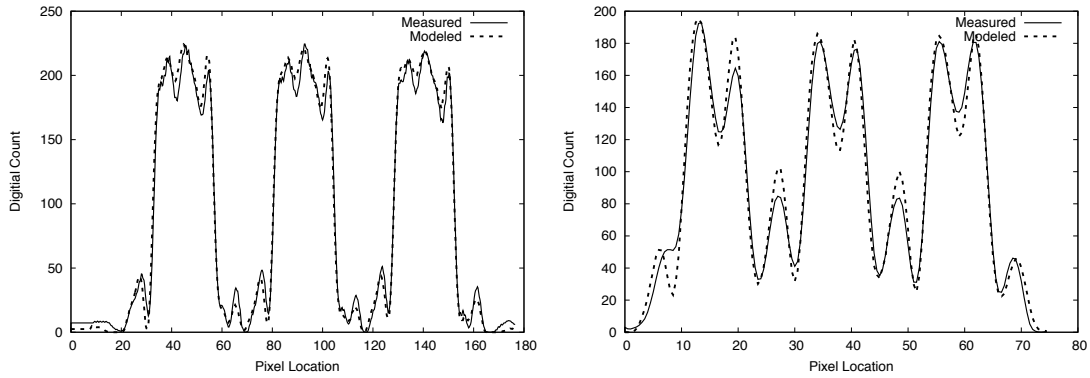
Figure 6.13: Enlarged version of bars 12-23 in the wavefront artifact validation experiment. Bar group numbers are marked in the measured image.

From these profiles, the average point-to-point average L1 distance for the ith profile was computed as,

$$\epsilon_{L1}(i) = \frac{1}{N}\sum_{j=1}^{N} |\text{Measured}_i(j) - \text{Modeled}_i(j)| \tag{6.8}$$

These results were then plotted as a function of a tri-bar group. This plot is shown in Figure 6.15. In general, error increased as the spatial frequency of the tri-bar increased. While this was not a strong trend, it does make sense. Since MTF is lower at those frequencies, the inverse filter is boosting more, which will magnify any errors in the image modeling process. Additionally, as bars shrink, the pixel phasing issue becomes more significant, although the pixel phasing in this experiment was a fairly close match, as was shown by the uneven bar spacing in the smallest tri-bars. For the lower-frequency bars, error was around or below 10 digital counts, which was roughly 5% of the bar magnitude.

From these profiles, the artifacts peaks and areas were measured manually. The specifics of these measurements was shown in Figure 6.9. For this analysis, only artifacts located in the gaps between bars and outside the first and third bar were measured.



(a) Vertical Tri-bar 0

(b) Horizontal Tri-bar 7

Figure 6.14: Sample registered tri-bar profiles from wavefront artifact validation experiment.

This was done for a number of reasons. Artifacts occur inside the bars, however, the artifacts from both sides of the bar tended to interfere with each other, making it difficult to put bounds on the area measurement. Since there was rarely a constant region inside a single bar, it was difficult to define a baseline value from which peaks or areas could be defined. Since the image should go to zero between bars, all signal in between bars could be attributed to artifacting. The resolution of the system was sufficient that modulation of the tri-bars did not become an issue until around bar group 13, at which point artifacts were no longer appearing between bars.

Figure 6.17 shows a scatter plot of measured peak heights versus modeled peak heights. Ideally for a perfect model, these measurement should lie on the line $y = x$, which is overlaid on the plot. In the horizontal plot, this appears to be the case in general, with the exception that the smallest peaks tend to be overestimated by about 5-10 digital counts. The vertical tri-bar plots fits slightly less well, although it is still close to the expected line. Error in this data set can be relatively easily explained, however. As can be seen in Figure 6.17, artifacts on vertical bars tended to be located closer to the bars than those on horizontal bars. In fact, it was sometime the case that the artifact was primarily positioned near the edge or inside the tri-bar, which was observed in Figure 6.12 in bars 8-11. Note that this corresponds to the peak in the L1 error in Figure 6.15. When the artifact is located very close to the edge, structured noise inside the bar will have a larger effect

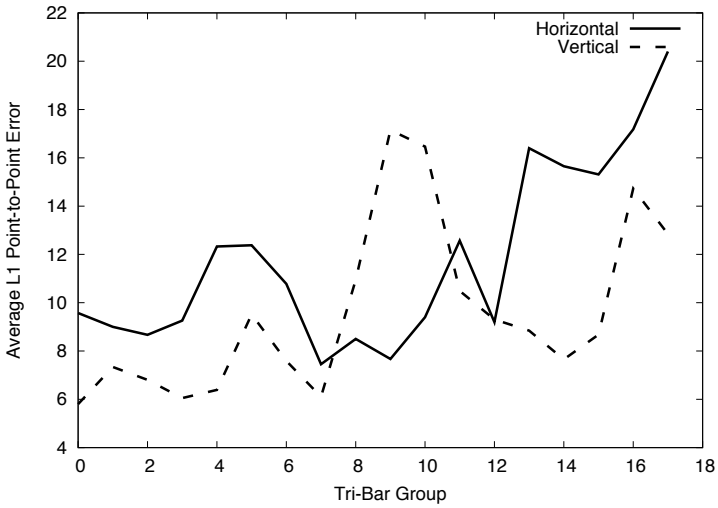
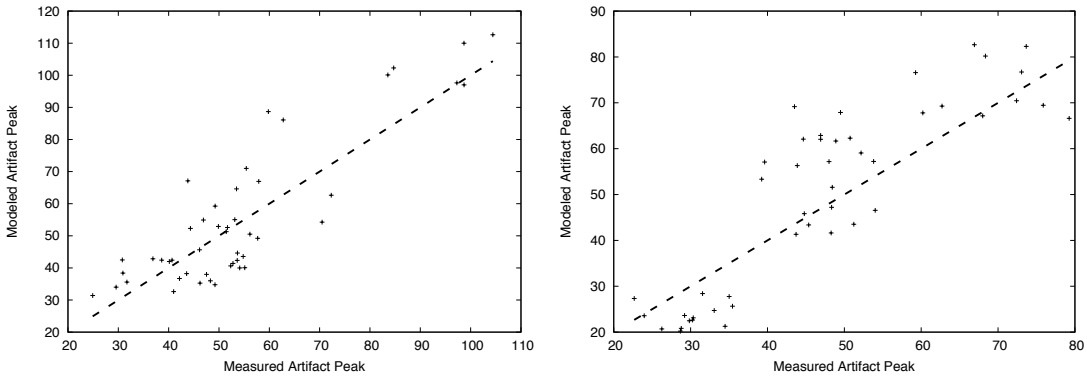


Figure 6.15: Average L1 error as a function of bar group for horizontal and vertical tri-bars.

on the artifact profile. This effect is to "break up" the artifact, which smoothes out the profile, lowering the peak and leading to an overestimation. In addition, it is possible this bias would have been balanced out if artifacts inside the bars were also analyzed.

In order to test how well the data fit the ideal $y = x$ line, which would imply a perfect model, a regression analysis was performed on the peak data. For this analysis, the bias in the fit was fixed to zero and a 95% confidence interval for the slope was computed. For the horizontal data, a slope of 1.024 was obtained, with a standard error of 0.029, leading to a 95% confidence interval of 0.966 - 1.081. For the vertical data, a slope of 1.064 was obtained, with a standard error of 0.030, leading to a 95% confidence interval of 1.005 - 1.124. These results match with the visual interpretation, as the horizontal tri-bar results matched well with the model, while the model tended to overestimate the artifacts for the vertical tri-bars. Again, since the artifacts on the vertical tri-bars were located close to the edges, any errors in the model, such as structured noise or wavefront error predictions, could have a large effect on the peak height or area.

Finally, Figure 6.17 shows similar scatter plots for the artifact area. As with the peak height plots, the horizontal data is well-centered around the $y = x$ line, while the vertical artifacts tend to underestimated at low magnitude and overestimated at high-frequencies.

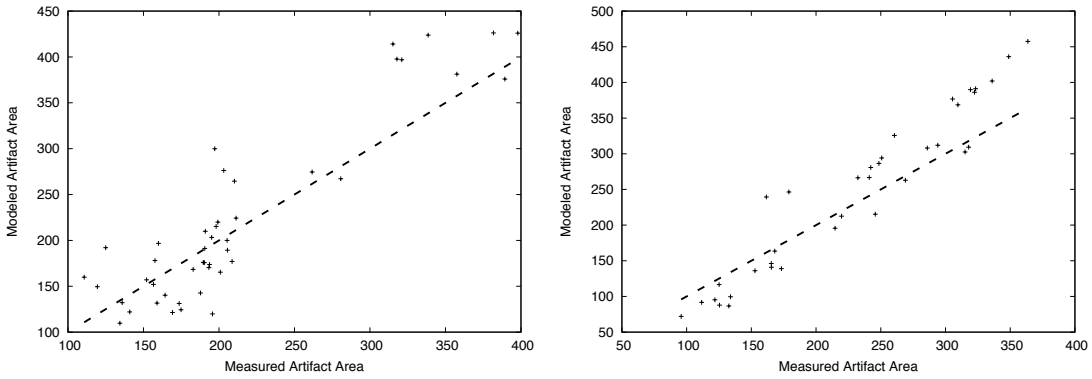


(a) Horizontal tri-bars (b) Vertical tri-bars

Figure 6.16: Artifact peak heights in measured and modeled data sets, overlaid with a 1-to-1 line.

For the horizontal tri-bars, the regressed slope was 1.059 with a 95% confidence interval of 1.003 - 1.114. For the vertical tri-bars, the regressed slope was 1.107 with a 95% confidence interval of 1.053 - 1.160. The area results agreed with the peak height results, although the differences from unit slope are higher, by approximately a factor of 2. Looking at Figure 6.14, artifacts tend to have a width in the neighborhood of 5 pixels, which was consistent across tri-bar groups. Since artifacts are not rectangular, the derivative of peak area with respect to height will be less than 5. In a simple simulation with Gaussian peaks, this derivative was computed as 1.86, so the factor of 2 is reasonable, since artifact shape varied from profile to profile.

While the structured noise and proximity of artifacts to the edges on vertical bars contribute to the error, there are also some simpler sources of error. As was mentioned earlier, the inherent wavefront error of the system was not characterized due to a lack of the necessary equipment. As such, the modeled and measured raw imagery were likely produced under slightly different circumstances. This can account for some of the difference in the artifact instantiation. The decentration of coma had a noticeable effect in Figure 6.12, as the streaks on the number 3 are at a slightly different angle. Finally, the input image was manually scaled and rotated to match the measured image. Any errors in this process or, more importantly, lens distortion will result in a difference in the predicted artifacts.



(a) Horizontal tri-bars (b) Vertical tri-bars

Figure 6.17: Artifact area in measured and modeled data sets, overlaid with a 1-to-1 line.

6.5 Spectral Bandpass Selection

6.5.1 Background

After examining the effects of wavefront error, the other main source of artifacts is error in the spectral weighting function of the system. When designing a sparse aperture system, signal-to-noise ratio is often a prime concern, due to the low fill factor. As [Breckinridge et al., 2008] and [Fiete et al., 2002] showed, integration times for sparse aperture systems can rise drastically with respect to the fill factor. Wide bandpasses can be used to slightly mitigate the signal-to-noise issues. However, large bandpasses also introduce the potential for error in the spectral weighting function of the system. Since the spectral response of the system is used when performing restoration, there is an implicit gray-world assumption introduced.

In order to minimize the effect of this gray-world assumption, a multi-band sparse aperture system could potentially be utilized. In this design, multiple bandpasses would be introduced, each with a smaller bandpass than the broadband system. Each bandpass would then be restored independently and then summed to restore the signal-to-noise ratio of the broadband system. The rationale behind this design would be that the gray-world assumption would be violated less in each individual bandpass, thus mitigating artifacting due to error in the spectral weighting function. This effect is illustrated in Figure 6.18. In this figure, a typical vegetation reflection spectrum is shown. Clearly, the spectrum

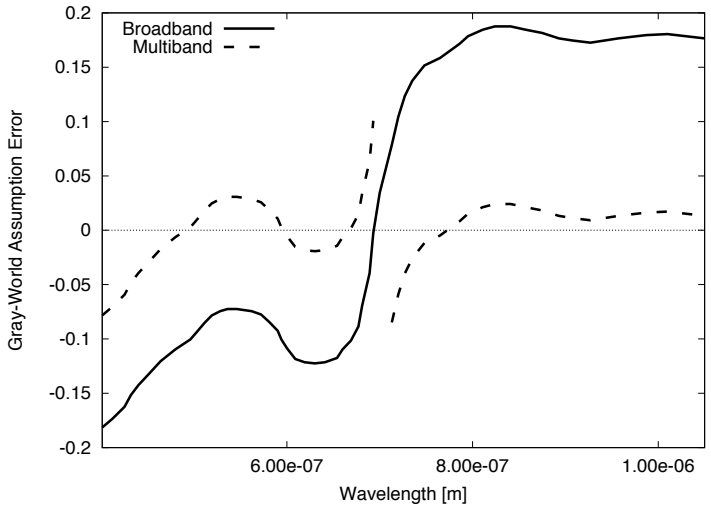


Figure 6.18: Splitting a system's bandpasses can decrease the error of the gray-world assumption in each band

violates the gray-world assumption over the shown bandpass. However, if the system's bandpass is split at 700 nanometers, then the gray-world assumption is much more valid on each side of the split.

While splitting bandpasses can theoretically reduce artifacting, the design choice is not without its drawbacks. As was described in Section 3.7, inverse filtering works better on data with low noise levels. In a photon-noise dominated system, splitting the bandpass reduces the signal-to-noise ratio in each band, thus reducing the effectiveness of inverse filtering in those bands. In addition, by narrowing the bandpass, less MTF smoothing occurs due to the spectral scaling of the optics OTF, which was illustrated in Figure 4.4. This results in less consistent performance over the frequency range of the system. Finally, splitting the bandpass can lead to registration issues between the bands, which can harm high-frequency performance in the recombined final image with ghosting artifacts.

6.5.2 Methodology

In order to create a demonstration of this multi-band system design, a number of spectral filters were introduced into the system. The first was a dual-band filter, which was attached to the light source. This dual-band filter was acquired from Edmund Optics

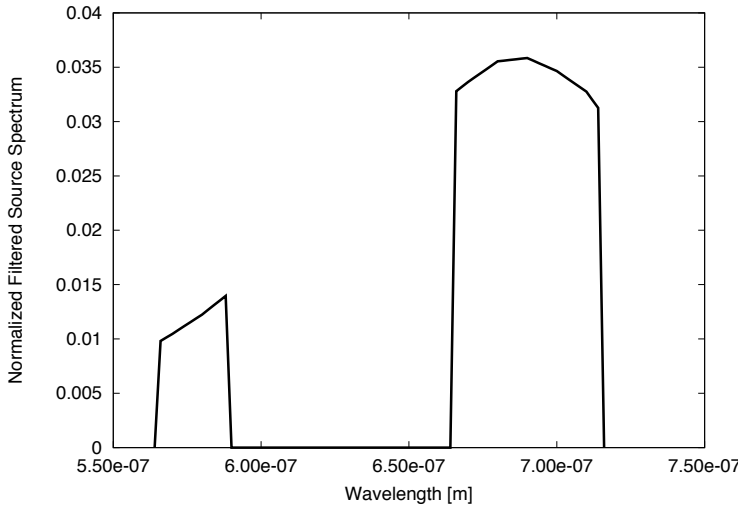


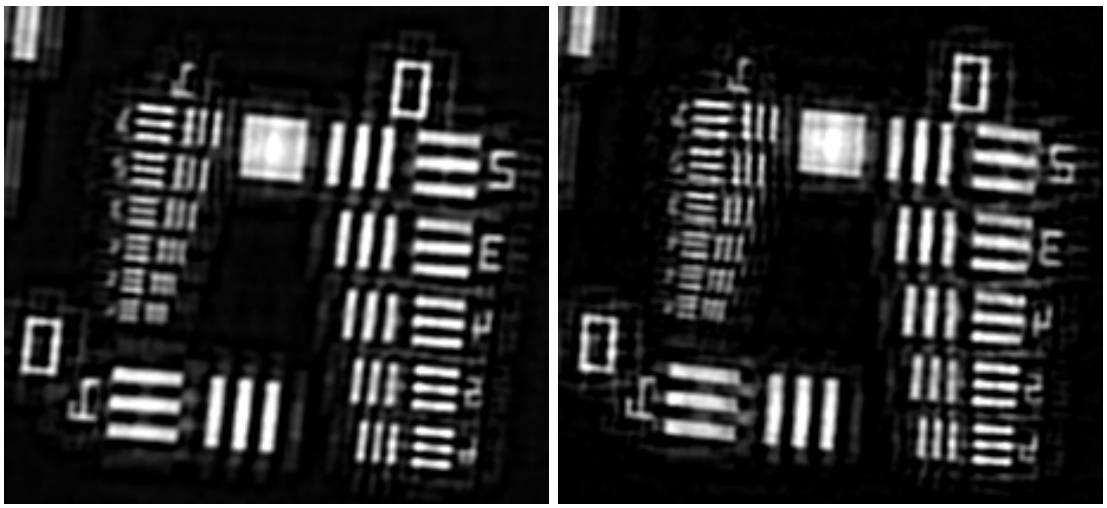
Figure 6.19: Illumination spectrum of laboratory system with 577/690 dual-band filter

and had approximately 98% transmission in both bandpasses, 565-589 nanometers and 665-715 nanometers, with approximately zero transmission elsewhere. The product of the filter transmission spectrum with the light source spectrum is shown in Figure 6.19.

Once inserted, the dual-band filter creates a spectral weighting function with high-frequency information. When imagery is restored without knowledge of this spectral weighting function, it is expected that artifacting will occur. To test this hypothesis, the USAF-1951 target from the earlier artifact validation experiment will be illuminated by the filtered spectrum and imaged under two different scenarios: broadband and multi-band. In the broadband scenario, the entire bandpass of the SBIG-8300M detector will be used. In this case, the quantum efficiency spectrum of the detector will be used as the spectral weighting function for the system during restoration. The multi-band scenario will introduce more filters into the system in front of the camera, a 650 nm shortpass filter and a 650 nm longpass filter. An exposure will be taken using each filter to simulate a multi-band sensor. The products of the detector quantum efficiency spectrum and the filter spectrums will be used as the spectral weighting functions during restoration of each band. After independent restoration, the restored images from each band will be summed to produce the multi-band image. The broadband and multi-band images will then be compared with respect to modulation performance and artifacting.

6.5.3 Results

6.5.3.1 Broadband Results



(a) Measured restored image (b) Modeled restored image

Figure 6.20: Enlarged version of bars 12-23 in the USAF-1951 target imaged under the broadband scenario.

Figure 6.20 shows the central portion of the restored USAF-1951 target for the broadband configuration, along with a prediction from the model. As can be seen, the results again match fairly well between the modeled and measured data sets. There are a couple of noticeable mismatches in this data set. The artifacts above the tri-bars on the right side of the image are oriented at slightly different angles between the modeled and measured images. This effect arises because of the decentration between the sparse aperture imaging system and the collimator. This decentration results in a mismatch between the coma predictions of the model and the actual coma present in the system, thus changing the orientation of smear artifacts. Unfortunately, since both the centration and magnitude of the coma are unknown, it was impractical to run an optimization to match the coma present in the measured images. It should be noted that the spectral and wavefront error data sets were collected months apart, which is why coma centration and pixel phasing differ between the data sets.

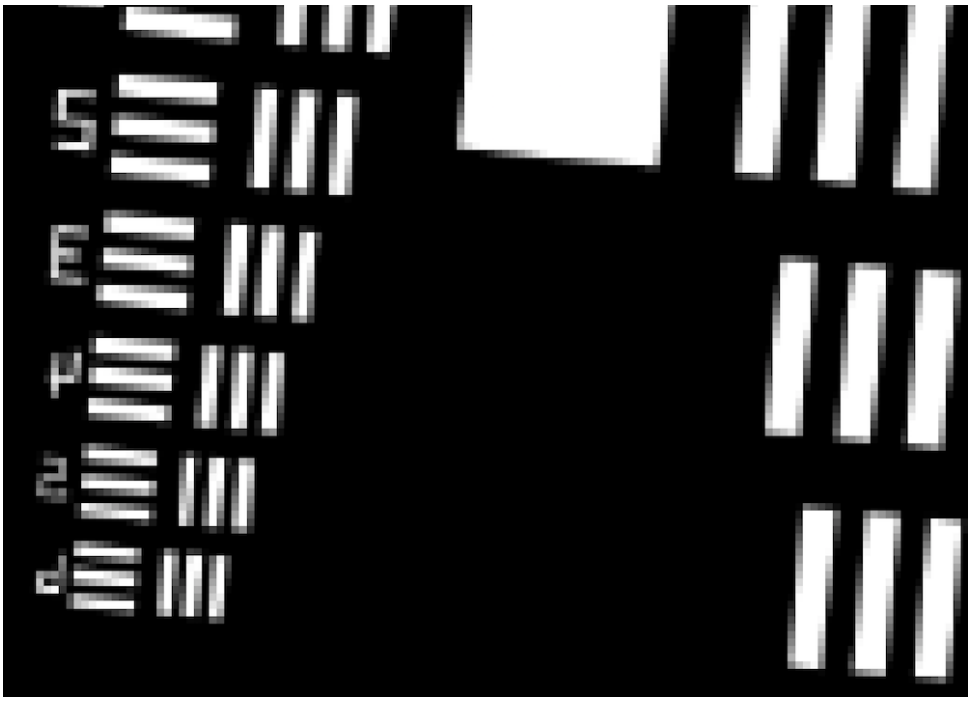


Figure 6.21: Enlarged version of the center of the input image to the model. The effects of pixel phasing can clearly be seen on the smallest bars.

In addition to the mismatch in coma artifacts, the smaller tri-bars on the left side have different contrast and spacing between the modeled and measured images. In the measured imagery, tri-bars have lower contrast and the three bars in a group are evenly spaced. This differs with the modeled data, where bar spacing is altered, which results in higher contract. In the wavefront error results, the uneven spacing and higher contrast were also present in the measured data, but they are not present in this broadband data set. This is simply a function of pixel phasing. Figure 6.21 shows the raw input image that is given to the model. As can be seen, the edges of the bars have gray values in between black and white due to anti-aliasing. While this precise instantiation of pixel phasing can occur in real data, as it roughly did with the wavefront error data set, it is not guaranteed. As can be seen in the figure, the number of anti-aliased pixels can approach or even outnumber the pure white pixels on the last series of tri-bars. Since this effect can change between the modeled and measured scenario, it is difficult to have high confidence in any conclusions drawn from these tri-bars, which is why they have been omitted from consideration and

analysis has been confined to tri-bar groups 0-17. In the future, it would be useful to have the ability to change pixel phasing in the model, however, this would likely require a multi-scale input image, as sufficiently increasing the resolution of the input image would have a large effect on model runtime.

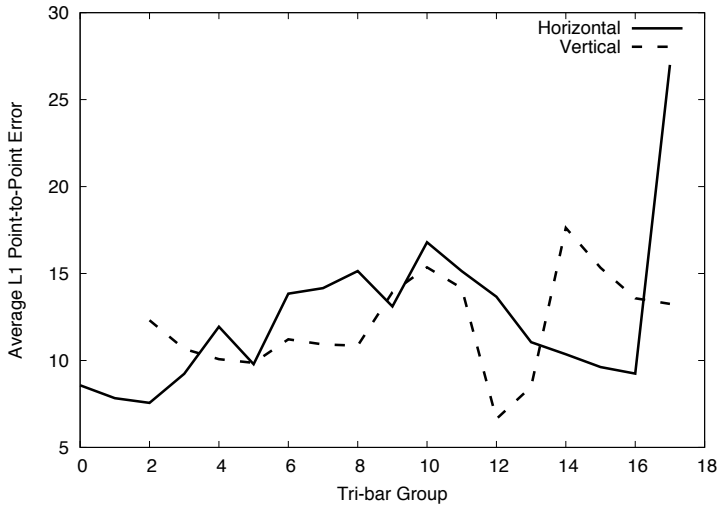
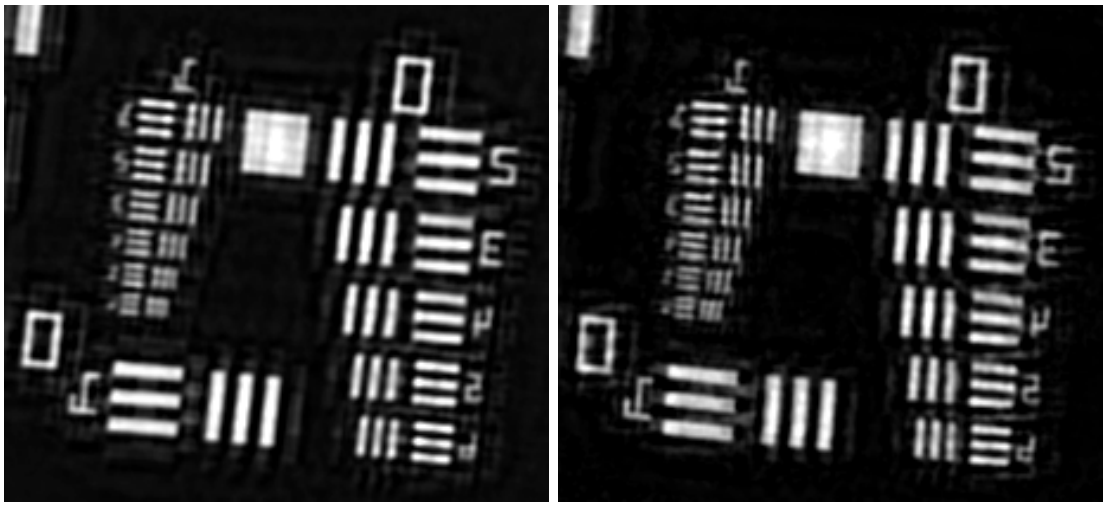


Figure 6.22: Average L1 error as a function of bar group for horizontal and vertical tri-bars illuminated by the spectrum in Figure 6.19.

As with the wavefront error data set, the average L1 error was computed for the broadband data set. The L1 error as a function of bar-group is plotted in Figure 6.22. The error in this data set was slightly higher than the error in the wavefront error data set. Averaged over all tri-bar groups (disregarding the one outlier), the error in the broadband data set was 11.86, compared to 10.68 for the wavefront error data set. This increase in error can be attributed to a couple of factors. While the effect is most notable on the smaller tri-bars, which are omitted from the L1 calculations, pixel phasing has an effect on all tri-bars and the pixel phasing instantiation given to the model was a closer match on the wavefront error data set. In addition, the lens was placed at the position of best focus for this collection. As such, the uncertainty in the wavefront error is relatively higher in this data set, since the dominant wavefront error present in this collection comes from the collimator, which was not characterized, instead of the imaging lens, which was characterized.

6.5.3.2 Multi-band Results



(a) Measured restored image (b) Modeled restored image

Figure 6.23: Enlarged version of bars 12-23 in the USAF-1951 target imaged under the multi-band scenario.

After examination of the broadband results, the same procedures were applied to the case of a multi-band system. One image was taken through a 650 nanometer shortpass filter and one image was taken through a 650 nm longpass filter. This cutoff wavelength was chosen because it fell in the middle of the two bandpasses on the dual-band filter, which was placed in front of the light source. Due to the limited dynamic range of the system, in order to get enough exposure for the shorter bandpass, a ND filter also had to be inserted with the longpass filter. This does change the spectral weighting function of the system. Since the wavefront error is constant between the two imaging scenarios, artifacting will be worse in the shorter bandpass, since wavefront error is defined relative to the wavelength of light. Thus, the multi-band system will start at a disadvantage relative to the broadband system, but will still exhibit better artifact performance. The insertion of this ND filter does not affect the comparison to the model predictions, however, since the ND filter can also be inserted into the model.

Figure 6.23 shows the central tri-bar pattern from the measured and modeled multi-band system. The results are similar to those in the broadband system, with overall

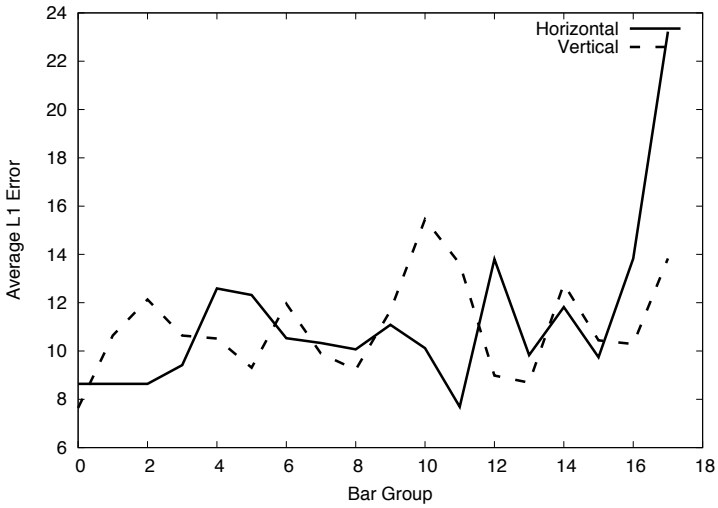


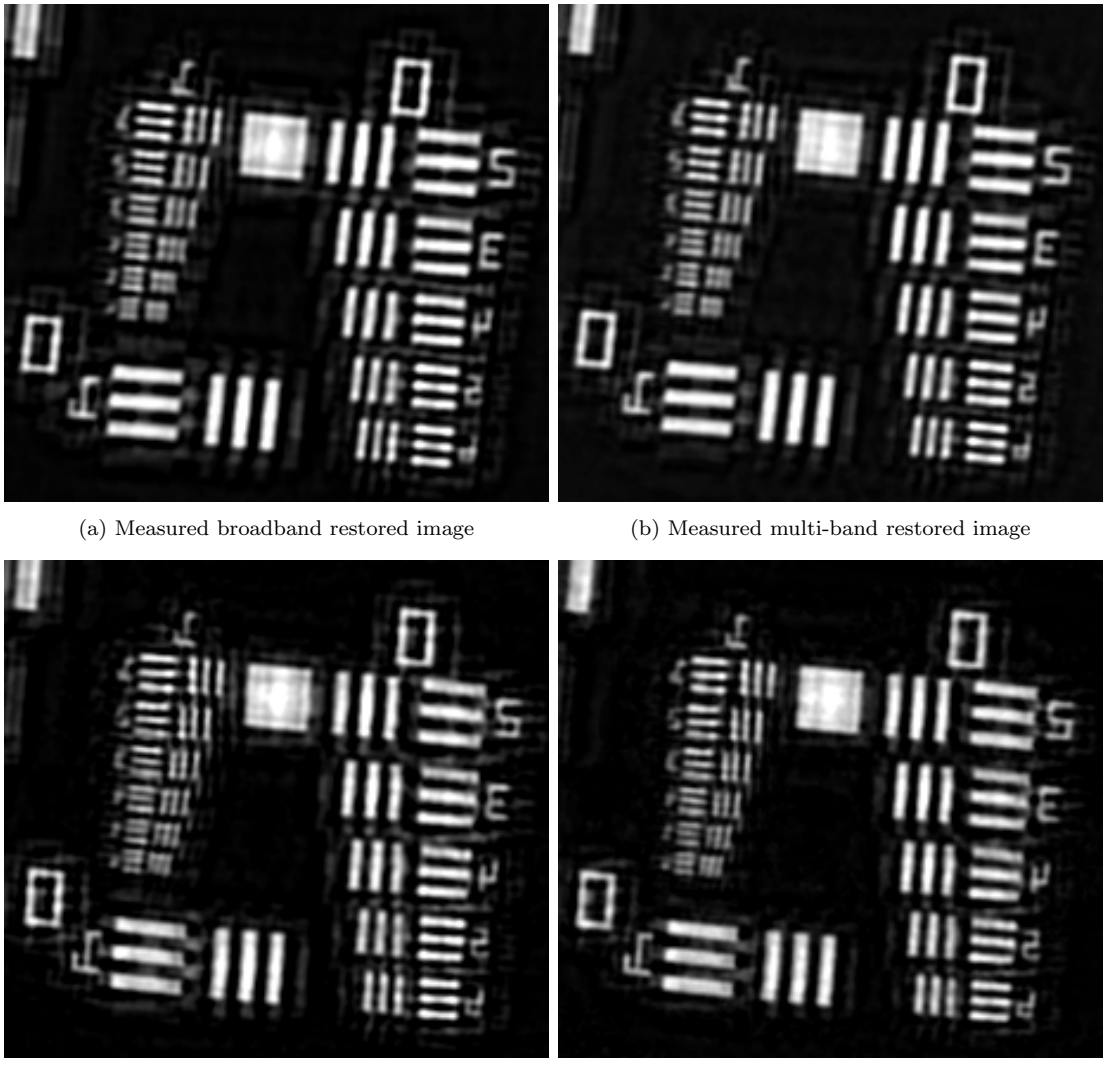
Figure 6.24: Average L1 error as a function of bar group for horizontal and vertical tri-bars for the multi-band system.

artifacting levels being comparable between the two images, with the exact instantiations differing slightly due to difficulties in modeling the exact wavefront error present in the system, which is most noticeable in the orientation of the smearing artifacts due to coma. As expected, artifacting is less severe in the multi-band system than in the broadband system, despite the wavefront error being constant between the two cases. The average L1 point-to-point error is plotted as a function of bar group in Figure 6.24. Over all tri-bars, the average L1 error was 10.89 digital counts, which is slightly lower than that of the broadband results and comparable to the wavefront error results. Again, the last tri-bar group exhibited considerable higher error as pixel phasing started affecting the modulation of the tri-bar and error was no longer dominated by artifacting differences.

6.5.3.3 System Comparison

Given the results obtained in the previous two sections, the broadband and multi-band systems can now be compared, along with the model predictions of these two systems. It is important that the gains predicted by the model are supported by the measured data, as this will establish that the model is an adequate stand-in for real data in future trade studies. Figure 6.25 shows the central tri-bar pattern in all four imaging scenarios. It is clear from visual inspection that the multi-band system has higher image quality in both

scenarios, despite the disadvantage of the higher weighting of the shorter bandpass in the multi-band scenario due to the ND filter. This image quality improvement is due to two related factors, increased contrast in the tri-bars and decreased artifacting.



(a) Measured broadband restored image

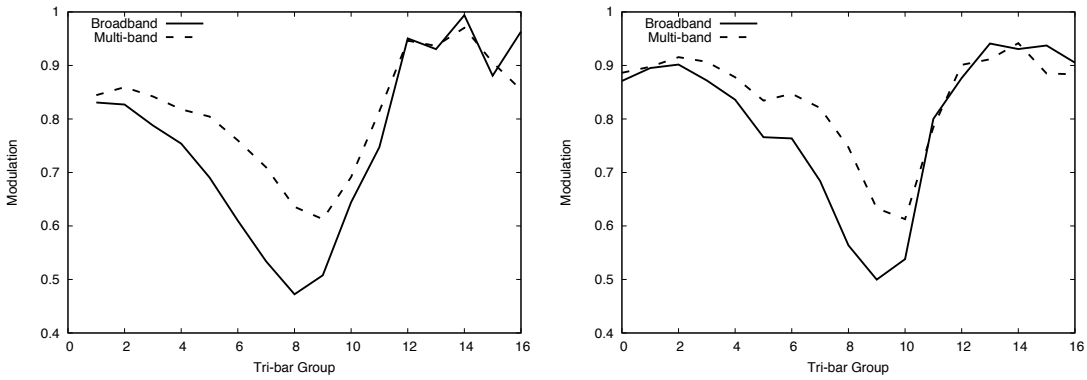
(b) Measured multi-band restored image

(c) Modeled broadband restored image

(d) Modeled multi-band restored image

Figure 6.25: A comparison of artifacting performance between the broadband and multi-band system in both real and modeled scenarios.

Splitting the bandpass has multiple effects on the post-processing operation. Decreasing the error in the gray-world assumption increases the accuracy of the "effective" OTF prediction, which decreases artifact levels and increases modulation performance in the restored image. However, the decrease in the signal-to-noise ratio from splitting the signal detracts from these performance gains. The signal-to-noise effect is most noticeable at frequencies with low OTF values, namely the highest frequencies or those between secondary peaks. As such, it is expected that modulation should increase on the low-frequency tri-bars and the tri-bars which share a spatial frequency with a secondary peak, while modulation should decrease on high-frequency tri-bars and tri-bars which fall in between secondary peaks. Artifacting levels should generally decrease.



(a) Measured data (b) Modeled data

Figure 6.26: Comparison of modulation performance between broadband and multi-band systems in both the real and modeled scenarios.

Modulation for a tri-bar can be quantified as,

$$M = \frac{DC_{\text{high}} - DC_{\text{low}}}{DC_{\text{high}} + DC_{\text{low}}} \tag{6.9}$$

where DC_{high} is the highest digital count in the bright regions of the tri-bars and DC_{low} is the lowest digital count between tri-bars. For traditional raw imagery, these peak values are sufficient, however, as was seen in Figure 6.14, tri-bars in restored sparse aperture imagery do not image as simple sinusoids. Instead averages will be taken over the peaks and troughs of the profiles. The bounds of these averages will be the first and last peaks

in the three bars and the first and last troughs in the two dark regions between bars. For higher-frequency tri-bars, profiles again begin to resemble sinusoids and the traditional extrema measurements will be used.

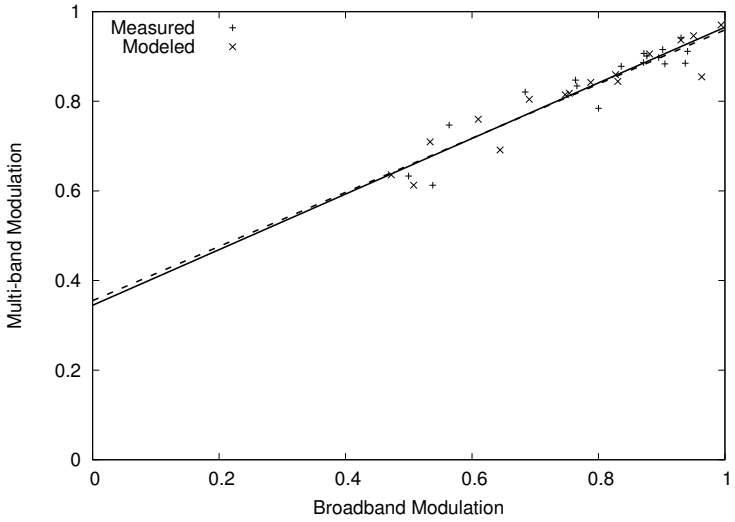


Figure 6.27: Multi-band modulation as a function of broadband modulation for both real and modeled data sets.

Modulation as a function of tri-bar group is plotted for both the measured and modeled data in Figure 6.26. This figure shows the same general trends in both measured and modeled data. Modulation starts out high and then decreases until tri-bar group 8 or 9. This decrease occurs due to artifacting, which increases the digital count in between bars. In this range, the multi-band system always outperforms the broadband system. Once tri-bars become small enough that artifacts no longer appear in between bars, modulation performance increases until image blur then causes it to again decrease, which occurs around tri-bar 13 in both measured and modeled data set. In this range the performance of the broadband and multi-band systems is very similar in both the measured and modeled data. In order to draw a direct comparison between the measured and modeled data, the broadband and multi-band modulations can be correlated and the regression lines can be compared. This analysis is shown in Figure 6.27. For the measured data, the regression

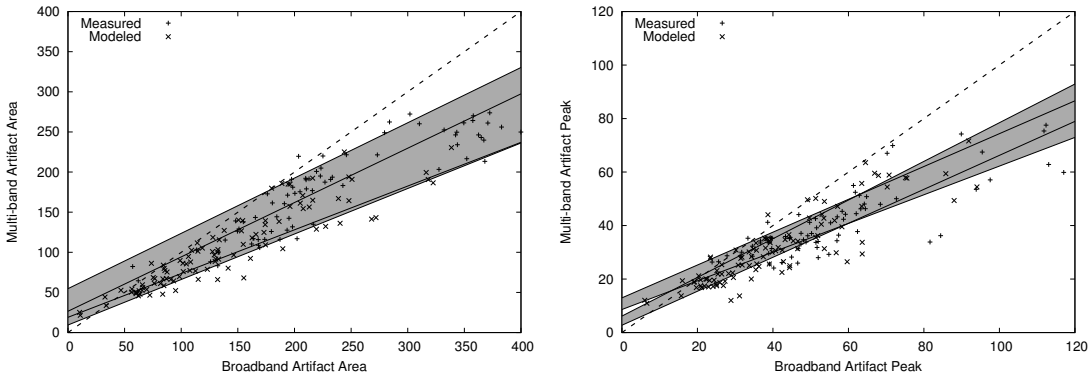
line (with its 95% confidence interval) is

$$M_{\mathrm{mb,meas}} = (0.620 \pm 0.129) \cdot M_{\mathrm{bb,meas}} + (0.345 \pm 0.104) \tag{6.10}$$

and the regression line for the modeled data is

$$M_{\mathrm{mb,mod}} = (0.604 \pm 0.115) \cdot M_{\mathrm{bb,mod}} + (0.355 \pm 0.089) \tag{6.11}$$

Clearly, from both the figure and the equations, the performance difference is consistent between the measured and modeled data sets, which means that the model is accurately predicting the difference in modulation performance between the broadband and multi-band system. The equations also give some insight into this performance difference. The multi-band system offers an improvement in modulation performance until a modulation of 0.8, at which point the broadband system has slightly better performance. Since the modulation does not exceed 0.8 until higher frequencies, where artifacts no longer appear between bars, this matches the behavior that was seen in Figure 6.26. Since the broadband image has higher signal-to-noise ratio, it should be able to be restored more effectively at the higher frequencies.



(a) Artifact area comparison

(b) Artifact height comparison

Figure 6.28: Comparison of artifacting metrics between the measured and modeled data sets.

In addition to improved modulation performance, the multi-band system also demonstrates less artifacting than the broadband system. As with the wavefront error data set, both artifact peak height and peak area were measured on both the measured and modeled

data sets. In order to draw a comparison between the measured and modeled data sets, the broadband and multi-band artifact metrics were correlated with each other for both data sets and linear regressions were performed. The results of these linear regressions are shown in Figure 6.28. The multi-band system consistently exhibited smaller artifact area and artifact peak height in both the measured and modeled data sets. For the measured data set, the regression line (with 95% confidence bounds) was given as

$$A_{\text{mb,meas}} = (0.617 \pm 0.072) \cdot A_{\text{bb,meas}} + (36.878 \pm 17.844) \tag{6.12}$$

For the modeled data set, the same regression line was

$$A_{\text{mb,mod}} = (0.622 \pm 0.055) \cdot A_{\text{bb,mod}} + (18.013 \pm 8.586) \tag{6.13}$$

As can be seen in the plot and the equation, the slopes of the two lines are essentially equivalent, while the modeled data set has a slight bias towards less artifact area. This can be caused by a number of sources, but a slight underestimation of wavefront error in the system would be the simplest explanation.

The peak height of artifacts was also analyzed in both sets and is also plotted in Figure 6.28. Of the two metrics, peak height exhibited more variability, which is logical, since it is more susceptible to be influenced by the structured noise present in the measured data. Regression analysis was again performed, however, since the data sets contained some outliers, the RANdom SAmple Consensus (RANSAC) algorithm [Fischler and Bolles, 1981] was used to pick the input points to the regression. Since not all point were used in the regression, the true 95% confidence interval is larger than the one produced by the regression analysis. For the measured data set, the regression line was

$$h_{\text{mb,meas}} = (0.575 \pm 0.039) \cdot h_{\text{bb,meas}} + (10.792 \pm 2.169) \tag{6.14}$$

and for the modeled data set, the regression line was

$$h_{\text{mb,mod}} = (0.679 \pm 0.044) \cdot h_{\text{bb,mod}} + (4.442 \pm 1.71) \tag{6.15}$$

As can be seen in the figure, these two confidence regions cross in the middle of the data set, thus making the slope the meaningful difference between these two fits. The

modeled data set has the steeper slope, indicating that artifact heights in the modeled multi-band data were generally higher than they were in the measured multi-band data. This could be explained by a couple of factors. While structured noise and error in the wavefront error estimates are again factors here, there is also the issue of registration. In the model, the two bands are perfectly registered and thus artifacts in the two bands will likely lie perfectly on top of each other, causing peaks to add together. If there is a slight, pixel-level registration error in the measured data, it could cause a slight mis-registration of artifacts, causing a decrease in peak height while not significantly affecting the artifact area. However, given that the confidence bounds are understated due to the use of RANSAC, the slopes are in reality fairly close to each other and this is a minor effect.

6.6 Conclusions

In this chapter, post-processing artifacts in restored sparse aperture imagery were examined. Simulations from the image quality model from Chapter 4 and previous work have indicated that these artifacts appear in imagery due to errors in the post-processing filter, which arise from errors in the wavefront error estimates or spectral weighting function. As such, experiments were performed to examine the effects of both error sources. A USAF-1951 target was inserted into the collimator and imaged with the laboratory system from Chapter 5. Two studies were performed to examine the two theorized sources of artifacting.

In order to examine the effects of wavefront error, the inherent wavefront error of the laboratory system and additional lens defocus were present in the imaging process, but not accounted for during restoration. These conditions were approximated and reproduced in the model. Point-to-point error, as well as artifact area and peak height were measured and compared to the model predictions. The average L1 error was approximately 5% of the magnitude of the signal level and artifacts predicted by the model were generally close in both shape, size and position to those seen in the modeled imagery. The peak height and area between the measured and modeled data were also well correlated. Sources of error were small, but explainable.

Errors in the spectral weighting function were examined in a slightly different manner.

Since there was inherent wavefront error in the system, it would have been impossible to separate artifacts due to wavefront error from those caused by spectral weighting error. Since in a realistic scenario, the spectral nature of the incoming radiance will be unknown, restoration in this experiment was performed only using the spectral response of the imaging system, thus introducing a gray-world assumption, while a dual-band filter was inserted into the system to create a highly spectrally-varying object. Images were then taken using the entire bandpass, for maximum spectral weighting error, and using a split bandpass to mitigate spectral weighting error. The split bandpasses were restored independently and then added to restore the original signal-to-noise level. The multi-band system demonstrated a measurable improvement in artifact performance, confirming the hypothesis that artifacts arise due to spectral weighting error. In addition, the model predicted this improvement with reasonable accuracy, meaning that the model could be useful in spectral band selection trade studies. Error in this study was slightly higher, but were explainable and a number of improvements to the model were suggested.

While the study presented in this chapter confirmed that predicted artifacts do in fact appear in real imagery and are accurate, there are still a large number of areas for future work. The largest issue is that while artifacts can be predicted, their effect on image quality is not yet understood. It seems likely that these artifacts have a more severe effect than the edge overshoot artifacts found in traditional imagery. Performing image quality studies, similar to those in [Garma, 2015], on simulated sparse aperture imagery would be a reasonable initial approach. It would also be useful to perform studies to see the effect of these artifacts on various automatic image analysis algorithms. Parallel to these types of studies is the need to quantify these artifacts. In this chapter, the area and peak height of these artifacts were used as metrics. It would be useful for a future study to correlate these metrics, or potentially others, with the results of image quality studies. In terms of artifact modeling, future improvements could be made to the model to handle the effects of structured noise and pixel phasing, which appear to be the two effects that cause most of the differences between modeled and measured data. Finally, the studies performed in this chapter could be reproduced in a setup where the wavefront error is more controllable. Given the constraints of the available equipment, wavefront error estimates had to be manually adjusted. A more sophisticated setup could examine the effects on artifacts of different forms of wavefront errors.

Chapter 7

Aperture Layout Optimization

7.1 Introduction

Previous research in sparse aperture imaging has typically been confined to a small set of known existing designs that have been used or theorized in the field of radio astronomy [Fiete et al., 2002] or have simple geometric properties [Meinel et al., 1983]. These designs are typically examined for their desirable frequency response properties or practicality for construction, but may not be optimal for every application. In this chapter, a design framework for sparse aperture systems, based on genetic algorithms, will be introduced. This framework can be used to discover optimal sparse aperture layouts for a given set of mission parameters. For example, this framework might be used to design an aperture pattern that maximizes perceived image quality, in order to produce imagery for use in a consumer product, such as Google Earth. Another use case might be to optimize an aperture that has maximal performance in a certain frequency range to match the size of a target that needs to be detected by the system.

The approach presented here is inspired by the successful previous usage of genetic algorithms in related problems. Genetic algorithms have been used in sparse antenna array design [Marcano and Durán, 2000], [Chen et al., 2006], [Chen et al., 2007], [Rattan et al., 2008] and sparse aperture imaging design for correlography [Henshaw and Guivens Jr, 1994]. While these works focus on different problems, they show that the usage of genetic algorithms on this type of design problem is feasible. This chapter will describe the adaptation of genetic algorithms to the problem of optimizing the layout of

sparse aperture telescope arrays for image quality in an incoherent imaging situation. A number of computational constraints from previous work will be relaxed and a few fitness functions for incoherent sparse aperture image quality will be presented. This work was published in [Salvaggio et al., 2016], but will be described in greater detail with an expanded background discussion here.

7.2 Background

Genetic algorithms are a set of biologically-inspired optimization algorithms that are based on the theory of evolution. They have a rich history in computer science tracing back to the 1960s. For an extended discussion of genetic algorithms, the reader is referred to [Mitchell, 1998]. This section will describe why genetic algorithms were selected for this problem and give a description of each step in the specific genetic algorithm used in this research.

When optimizing sparse aperture layouts, MTF properties and practicality for construction are generally of prime concern. More traditional optimization algorithms, such as Gauss-Newton or Levinburg-Marquardt, require derivatives of the optimization function with respect to the optimization parameters. Typically the parameterization of a sparse aperture consists of position and shape of the subapertures, making derivatives expensive or impossible to compute. Given that derivatives are unavailable, there are two options for optimizations: gradient-based methods with numerical derivatives or derivative-free optimizations. Gradient-based approaches normally require a starting estimate as initialization. If this estimate is not sufficiently close to the global optimum, there is a large risk of convergence to a local extrema in the cost or fitness function. Obtaining such an estimate is not generally possible in the sparse aperture layout problem, leaving derivative-free optimization as the most suitable option.

Genetic algorithms are one of many derivative-free optimization techniques that utilize randomness instead of derivatives. Particle swarm optimization is another notable example that would be applicable to this problem. Both of these algorithms consider a number of candidate solutions, which increases robustness to local extrema, giving them advantages over other derivative-free optimization algorithms, such as Nelder-Meade, for

this problem. Genetic algorithms also excel on problems that are compactly parametrizable, which is true of sparse aperture telescopes with discrete subapertures, especially if design constraints are enforced through the parameterizations.

While genetic algorithms offer some advantages over gradient-based optimization approaches, the reliance on randomness does come with some drawbacks. Termination of the optimization is an instance of the halting problem, as there is no guarantee that a better solution does not exist. The precision of the gradient-based algorithms is also hard to reproduce without greatly increasing the run-time. As such, this work will take the approach of tuning genetic algorithms for both global search and progressive local refinement, resulting in multiple optimizations being run to obtain a precise result. This approach improves precision and runtime, while sacrificing robustness to local minima. This robustness can be regained, however, by running multiple global search algorithms and selecting the best output for refinement.

Given the selection of genetic algorithms for this problem, the role of each step of the algorithm will now be described. The implementation of these steps will be described in the Methods section. Like most optimization algorithms, genetic algorithms are an iterative process. An overview of one iteration is given in Figure 7.1.

As shown in the figure, during each iteration, there exists a set of potential aperture configurations, termed the *population*. During each iteration, a set number of new configurations are derived, or bred, from the population. This process involves two steps, *crossover* and *mutation*. Crossover combines attributes from the two parent configurations to create a new configuration. Mutation then introduces random variation into the newly-produced configuration. Finally, a new population is constructed from the union of the initial population and the newly-created configurations. Details on each step in the algorithm will be given in the following sections.

7.2.1 Crossover

The task of the crossover operator is to produce a new configuration from two existing members of the population. The philosophy behind this operation is that if two population members both have attributes that produce high values of the fitness function, then

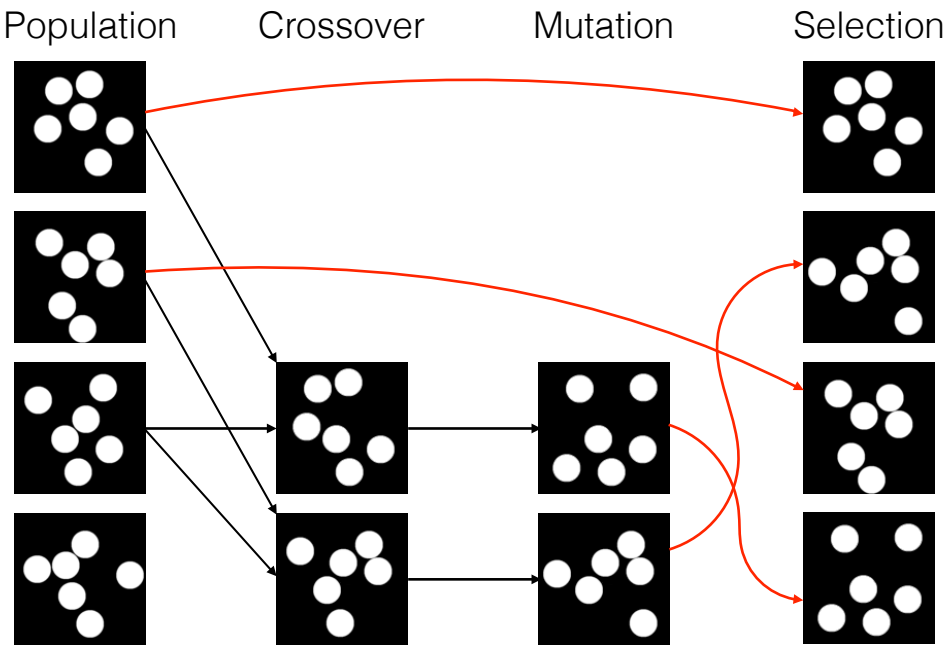


Figure 7.1: Overview of one iteration of a genetic algorithm optimizing aperture design

a better configuration might be found by combining the positive attributes of both. Thus, it would make sense to use the best members of the population as inputs to the crossover operation. However, repeatedly selecting the two best members of the population would be highly susceptible to local maxima in the fitness function. To achieve a balance between robustness and convergence speed, inputs to the crossover operator are selected using the fitness values of the population members as a probability density function.

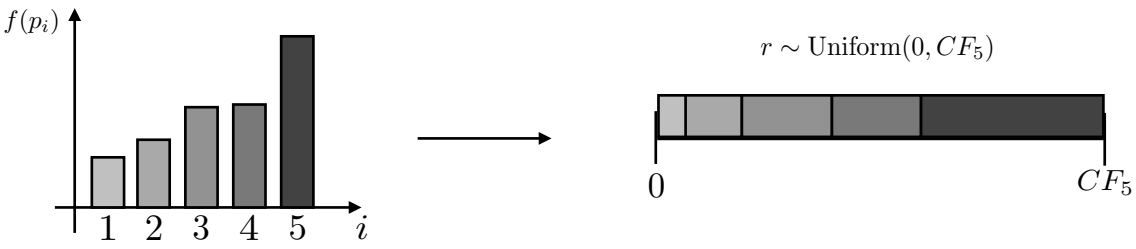


Figure 7.2: Illustration of crossover operand selection process

At the beginning of each iteration, the fitness function is evaluated for each population member. The population is then sorted in decreasing order of fitness and a cumulative fitness function is computed. Defined recursively,

$$CF_0 = 0 \tag{7.1}$$

$$CF_i = CF_{i-1} + f(p_i) - f(p_{\min}) + \epsilon \tag{7.2}$$

where $f()$ is the fitness function, p_i is the i-th population member, $p_{\min}$ is the population member with the smallest fitness function and ϵ is a small number to give a baseline probability of all population member being selected. The construction of the cumulative fitness function is visualized in Figure 7.2. Once the cumulative fitness function has been computed, members are selected randomly, weighted by the cumulative fitness function. A random number r is selected uniformly between 0 and CF_N. The selected member is the member i, where CF_i is the first value in the cumulative fitness function greater than r [Lipowski and Lipowska, 2011]. Crossover requires two operands and thus r is repeatedly generated until two different members are selected.

Since each crossover operation produces one new population member, the user is free to select how many new members the algorithm should create in each generation. Since there is overhead associated with each generation, creating more candidates can speed up computation and reduce the number of required iterations. However, generating more members per generation can have diminishing returns if there is a large amount of redundancy in the outputs of the crossover operator. The optimal number of configurations to create will depend on the implementation of the crossover operator. If the operator can create many new, non-redundant configurations, then a larger number of new configurations will be ideal. Otherwise, the parameter should be set to a lower value. In practice, this parameter is set by the user and must be empirically tuned.

7.2.2 Mutation

Like crossover, mutation is a biologically-inspired operator intended to explore the search space. Whereas crossover combined attributes from members already in the population, mutation introduces new attributes, or modifies existing attributes in new ways. This operator is important to drive the genetic algorithm into unexplored portions of the search

space. Unlike crossover, mutation is a unary operator, meaning it operates on one population member, potentially changing its parameters. Unlike crossover, the mutation operator is deterministically invoked. That is, every output of the crossover operator is passed to the mutation operator. The mutation operator will then internally use randomness to introduce (or not introduce) its variations.

As was mentioned earlier, this research will utilize a two-stage optimization approach: global search and local refinement. Mutation is used in both steps but is the only operator used in local refinement. If the population is initialized as copies of a starting guess and an identity crossover operator is used, then a mutation operator that only introduces small changes can explore the local neighborhood around the initial guess. Evaluating the fitness function and using the selection process detailed in the next section can make the genetic algorithm function in a similar manner to a local heuristic search, such as the A* algorithm [Hart et al., 1968].

7.2.3 Selection

One of the key parameters for a genetic algorithm is the number, M, of new population members to create during each iteration. For each new member, the crossover operation is run and the mutation operator is then run on its output. These two operators are not guaranteed to result in a valid population member. So, the fitness function is evaluated for the new member to determine validity. If the new member is invalid, the process is repeated until a valid new member is produced. Once this process is finished, there will exist M new population members, along with the existing population of size N.

The goal of selection is to construct a new population of size N from the new group of $N+M$ population members. The goal of each iteration is to move the population closer to the global optimal solution. As such, it would be expected that the best N members should be taken as new population. However, the same argument applied to crossover operand selection also applies here, as does the solution. Instead of deterministically selecting the N best members, the cumulative fitness function is recomputed, by Equation 7.1 and 7.2. The same random selection process is then repeated until N unique members have been chosen to form the new population. Once this process completes, the next iteration begins, using the new population.

7.2.4 Termination

Termination of a genetic algorithm can be a difficult task. The core issue is that termination is an instance of the halting problem [Turing, 1937]. That is, given the fact that we are randomly exploring a continuous space, there is no guarantee that the optimal solution will ever be found. Furthermore, given the current optimal solution, there is no way of determining whether a better solution exists. As such, the user must decide what entails an acceptable solution for the given optimization.

To solve this issue, there are a large number of potential strategies. The simplest among these is to fix the number of iterations or execution time. Upon reaching either of these limits, the best member in the population is taken as the optimal solution. While these solutions are both simple, they offer no guarantees as to the quality of the solution. Another simple solution is to define a threshold for the fitness function. Once a solution exceeds this threshold, it is taken as "acceptably optimal" and the iteration is terminated. This solution gives guarantees on the quality of the solution, but not its optimality. If a well-defined specification for system performance is provided, then this is often the best strategy. However, if the fitness function is a weighted combination of multiple performance metrics, then computing the threshold can be challenging and the threshold loses its semantic meaning. Since the fitness function is a projection into a one-dimensional fitness space, there may be combinations of performance metrics that exceed the fitness threshold but do not satisfy the specification. In such a case, more sophistication is needed to apply thresholds to each component of the fitness function.

If the simpler methods are not sufficient, some more sophistication can be used. If there is an acceptable level of precision for the algorithm, the search space can be discretized. Once discretized, a local search can be assumed to have converged if the best solution does not change after a user-defined number of iterations. The user can also look at a plot of fitness function versus iteration number and determine a point of diminishing returns at which to stop the iteration in the future. However, none of these potential solutions guarantee a global optimal solution. Since this is an instance of the halting problem, such a solution does not exist and the user must decide which of the termination strategies is best suited for the given optimization.

7.3 Previous Work

This optimization framework builds directly on the sparse aperture modeling described in Chapter 4 and validated in Chapter 5. In addition to the previous work in sparse aperture image quality described in previous chapters, a few studies have looked at the design of sparse aperture telescopes. [Breckinridge et al., 2008] focused on the tradeoff between fill factor and image quality for several non-optimized designs. Another related study is presented in [Miller et al., 2007], which focuses on the Golay apertures, but takes an in depth look at optimizing both the number of apertures and the expansion factor. While [Salvaggio et al., 2015] did laboratory studies with broadband sources, [Miller et al., 2007] instead used monochromatic illumination. As a result, that study was able to explore the use of phase retrieval on a sparse aperture system, although the authors only found the technique to work under heavily constrained circumstances. This work looks to build directly on [Salvaggio et al., 2015] and [Miller et al., 2007] and provide a mechanism for finding alternatives to the Golay apertures. While the Golay apertures are optimal for frequency coverage, image quality is a complex issue and these apertures may not be the answer for all situations.

Due to their flexibility, genetic algorithms can be applied to a vast array of problems and have seen previous use in related problems. One such problem is that of sparse antenna array design [Marcano and Durán, 2000], [Chen et al., 2006], [Chen et al., 2007] and [Rattan et al., 2008]. The goal of these studies is to design an antenna array that produces a radiation pattern with a minimum side lobe level. While the research presented here will employ similar methods to these previous studies, the goals of the optimizations vary significantly. Image quality for sparse aperture systems is a more complicated issue than for conventional imaging systems and lacks a definitive metric. Finding such a metric is beyond the scope of this research, however, the results of some simple metrics will be demonstrated, as well as methods for applying construction constraints and examining the effects of subaperture shape.

A more directly-related previous work is presented in [Henshaw and Guivens Jr, 1994], where genetic algorithms were used to optimize a sparse aperture imaging array for correlography. In correlography, an object is illuminated by a laser and the measured speckle pattern is used to reconstruct an incoherent image. The quality of an aperture for cor-

relography in [Henshaw and Guivens Jr, 1994] was defined as the non-redundancy of the system MTF. Due to the computational constraints at the time, apertures were confined to a coarse grid and a 2D to 1D mapping function was required, which had noticeable effects on the result. As a result of the constraints, the MTF was a binary function, resolution was low and all peaks were identical and could not overlap. Unfortunately, using only a non-redundancy metric in incoherent imaging will lead to unacceptable levels of artifacting during image restoration, so new metrics will need to be developed.

In the proposed genetic apertures framework, all calculations are performed in two dimensions and are not quantized to a grid. This allows for more complicated fitness functions that incorporate the value of the MTF. This is a necessary improvement if genetic algorithms are to be used for optimizing image quality in an incoherent imaging situation. For example, in order to make a system more robust to post-processing artifacts, peaks may need to overlap in order to smooth the MTF and boost MTF values relative to the image noise. Such overlap is also necessary for low-contrast imaging scenarios, as was examined in [Breckinridge et al., 2008]. Additionally, the shape of the subapertures can be varied to control the shape of peaks. The gridded approach in [Henshaw and Guivens Jr, 1994] is also unable to account for the effects of a system's spectral bandpass, as it was intended for laser illumination. The approach in this research can be easily extended to support spectral effects (currently with a linear effect on runtime). This research also presents new fitness functions that focus on the image quality of sparse aperture imagery.

7.4 Methods

Figure 7.3 shows a graphical overview of the genetic apertures framework. In order to run an optimization, three components must be specified: the parameterization, fitness function and search strategy. The parameterization is the search space of the optimization and should be as low-dimensional as possible. The fitness function is the function that is optimized by the genetic algorithm. More desirable aperture layouts should result in higher values of the fitness function. Finally, the search strategy describes how the search space defined by the parameterization is explored. This consists of the initialization of the population and implementations of the crossover and mutation operators. Examples of each component will be given, along with some example results from several combinations

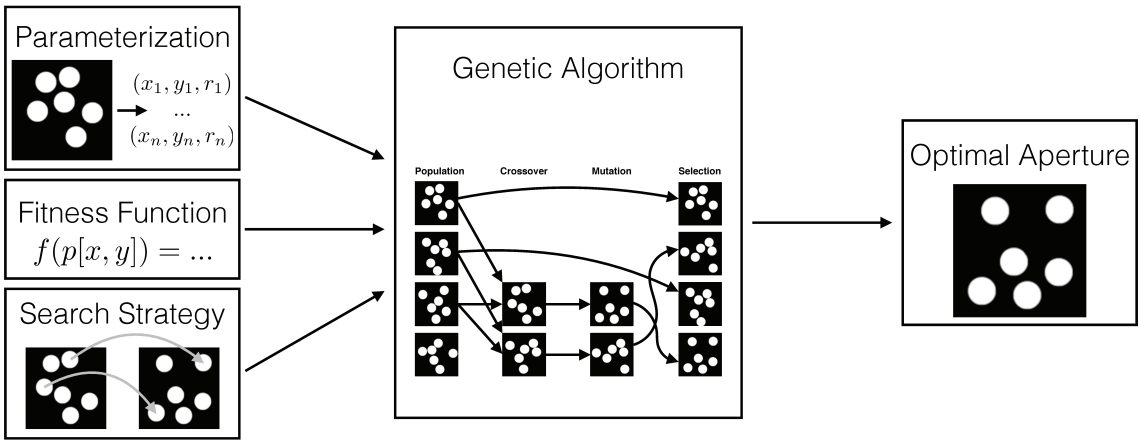


Figure 7.3: An overview of the genetic apertures framework. The boxes on the left must be specified by the user for each optimization, while the genetic algorithm stays constant.

of components. The software used in this work is written in C++ and is hosted at (`https://github.com/psalvaggio/multaptelescope`). All optimizations described in this research were run on a 2.6 GHz Intel Core i7 processor in under one hour, without parallelization.

7.4.1 Parameterization

As was mentioned previously, genetic algorithms excel on low-dimensional problems. As such, a compact parameterization of the sparse aperture layout is important to obtaining good performance out of the optimization framework. Using the parameterization to enforce design constraints is an ideal way to both lower the dimensionality of the optimization and ensure the optimal aperture will be practical for construction.

In this work, analysis was confined to sparse apertures composed of a discrete number of circular subapertures. Since circular optics are a common choice for framing remote sensing systems, this was a realistic design constraint for a demonstration of the system and also allowed for efficient approximation of the system MTF in fitness functions. As such, the most straightforward parameterization would be a list of (x_i, y_i, r_i) triplets for the position and radius of each subaperture.

A common design constraint in this parameterization might be to keep the subapertures inside some encircled radius, r_{enc} to bound the x and y dimensions. The r dimension in this parameterization also needs to be constrained so the optimization does not revert to a filled aperture. A common figure used in sparse aperture design is the fill factor, F, which is the percentage of the area inside r_{enc} that is filled with glass. This holds the sum, $\sum r_i^2$ to a constant value. Taking this one step further, if the component telescopes are already specified and only the layout is desired, the r dimension can be quantized to match available telescopes.

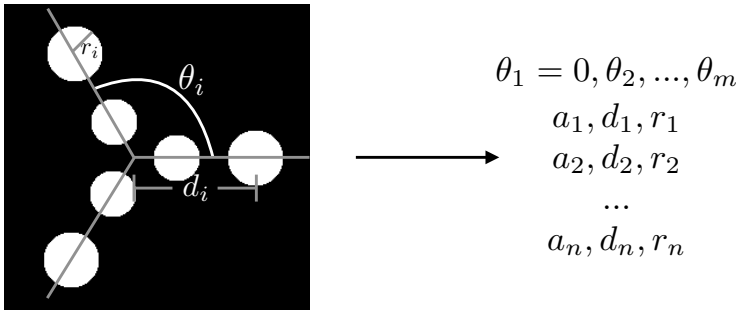


Figure 7.4: A parameterization for a sparse aperture array with m discrete arms. θ_i refers to the arm angle angle of the ith arm. For rotationally invariant fitness functions, θ_1 can be fixed to 0. a_j, d_j and r_j refer to the arm index, radial displacement and radius of the jth subaperture, respectively.

In addition to MTF characteristics, practicality for construction is another concern for sparse aperture layout design. While construction terms can be used in the fitness function, changing the parameterization can be an efficient way to shrink the search space and ensure a constructible layout. For example, in order for a space-based instrument to fit into a certain launch vehicle, perhaps all subapertures need to exist on a small number of foldable arms. Such a situation is shown in Figure 7.4. In this case, the parameterization would be the arm angles, θ_i and a triplet (a_j, d_j, r_j) for the arm index, offset and radius of each subaperture. Similar constraints can be placed on the r dimension and the arm indices are quantized, leading to a smaller search space.

7.4.2 Fitness Function

The role of the fitness function is both to ensure a candidate aperture is valid and to map that candidate aperture layout to a scalar value, indicating the quality of that layout. The genetic algorithm will attempt to maximize the value of this function, yielding the optimum aperture. Ideally, this function will be designed to have a sharp peak and no local maxima, although this is often difficult to achieve or verify in practice. In the case of sparse aperture optimization, the fitness function typically aims to achieve a certain property in the MTF of the system. In this study, three distinct fitness functions were used.

7.4.2.1 Golay Validation Study

In order to validate the genetic apertures framework, a verification of Golay's solution [Golay, 1971] was attempted. The goal was to reproduce the Golay-6 aperture pattern, shown in Figure 7.5 (b), as this is a common aperture seen in sparse aperture studies, such as [Miller et al., 2007] and [Fiete et al., 2002]. In Golay's work, he claims that his apertures maximize non-redundancy and compactness of the MTF. His claim is logical, given the hexagonal packing of MTF peaks.

In order to reproduce Golay's result in this framework, a fitness function had to be constructed. A direct mapping of Golay's language to a fitness function would be a weighted sum of non-redundancy and compactness terms.

$$f_{\text{Golay}} = NR(MTF(\xi,\eta)) + \gamma \cdot C(MTF(\xi,\eta)) \tag{7.3}$$

$$NR(MTF) = \iint \lceil MTF(\xi,\eta) - T \rceil \; d\xi \; d\eta \tag{7.4}$$

$$C(MTF) = \left(\sum_{i=1}^{N} \sum_{j=1}^{N} (x_i - x_j)^2 + (y_i - y_j)^2 \right)^{-1} \tag{7.5}$$

where γ is a relative weighting term, $\lceil \cdot \rceil$ is the ceiling function and T is an MTF threshold to consider a frequency resolved above the noise of the system. The ceiling function serves to round any MTF value over T up to 1. Another method would be to apply a

monotonically decreasing weighting function, W, to the MTF, such as

$$f_{\text{Golay}} = \iint W(\rho) \cdot \min(MTF(\rho, \phi), T) \; d\rho \; d\phi \tag{7.6}$$

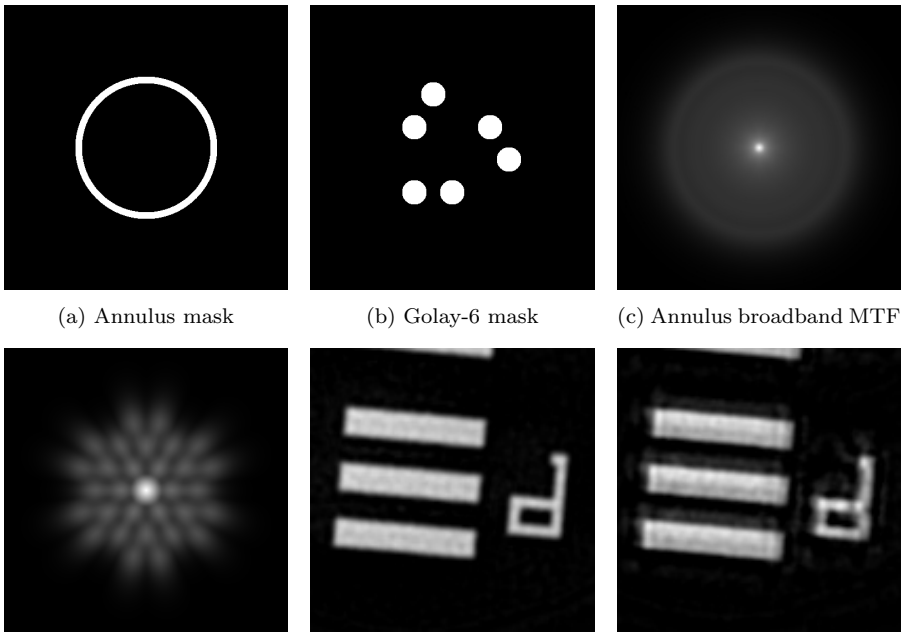
where ρ and ϕ are polar spatial frequency coordinates and T is again an MTF threshold, although this time it is used to bound the MTF support at a given frequency, encouraging the MTF to cover more frequencies. Simply setting the value of T to a small value over 0 will encourage maximum non-redundancy, while the monotonically decreasing property of W will encourage compactness. As will be shown in the Results section, both approaches reproduce Golay's results.

7.4.2.2 Discrete Annulus

Another common sparse aperture design seen in previous studies, such as [Fienup, 2000], is the annulus, or a circular aperture with a very large central obscuration. This design produces an MTF that is essentially constant, albeit low, out to the cutoff frequency. This is attractive in that using an annulus over a large bandpass only has an effect on the MTF at the highest frequencies, as can be seen by comparing the broadband MTF in Figure 7.5 (c) and the monochromatic MTF in Figure 7.6. A large drawback of sparse apertures are the image artifacts that occur due to post-processing restoration filters. The annulus design mitigates this drawback, as shown in Figure 7.5, but would be difficult to both construct and launch into space. The genetic apertures framework can be used to reproduce some of the positive aspects of the annulus in a constructible system, such as the discrete arm array from Figure 7.4.

In order to emulate the annulus in a discrete subaperture system, it must be described with a fitness function. While the Golay design exhibits maximum non-redundancy, it does so under the constraint of discrete sub-apertures. The annulus has support at every theoretically possible frequency for its diameter, and thus a non-redundancy term is also necessary here. However, the annulus is attractive due to the radial smoothness of its MTF, which is the source of its robustness to large bandpasses. As Figure 7.6 shows, this smoothness can be characterized by looking at enclosed MTF area as a function of radial frequency. The fitness function used to emulate an annulus in this work is given by

$$f_{\text{Annulus}} = \iint \lceil MTF(\xi, \eta) - T \rceil \; d\xi \; d\eta + \gamma \cdot r'_{\text{area}} \tag{7.7}$$



(a) Annulus mask (b) Golay-6 mask (c) Annulus broadband MTF

(d) Golay-6 broadband MTF (e) Simulated restored annulus image (f) Simulated restored Golay-6 image

Figure 7.5: Difference in post-processing artifacts between the annulus and Golay-6 sparse aperture designs. Notice the gradient in the bars and ghosting above each bar in (f). Both designs were given the same amount of coma and defocus wavefront error. Image simulation and restoration was performed with the modeling framework from [Salvaggio et al., 2015].

where the first term is the non-redundancy term from Equation 7.4 and r'_{area} is the correlation coefficient of the cumulative distribution from Figure 7.6 over a range of middle frequencies. Since most cumulative distributions will have high correlation coefficients, the range [0.98, 1] is linearly stretched to the range [0, 1] with everything under 0.98 being thresholded to 0.

7.4.2.3 Acutance

Acutance is a standard measurement of perceived image quality published by the IEEE Cell Phone Image Quality group [Baxter et al., 2012]. This metric takes into account the contrast sensitivity function of the human visual system, plotted in Figure 7.7 and

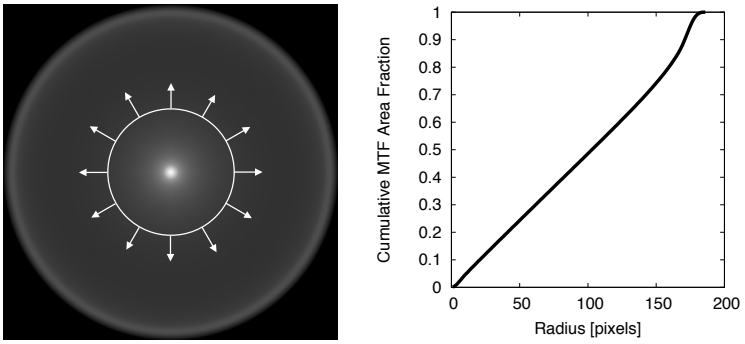


Figure 7.6: Enclosed MTF area as a function of radial frequency. The large linear region is characteristic of the annulus design.

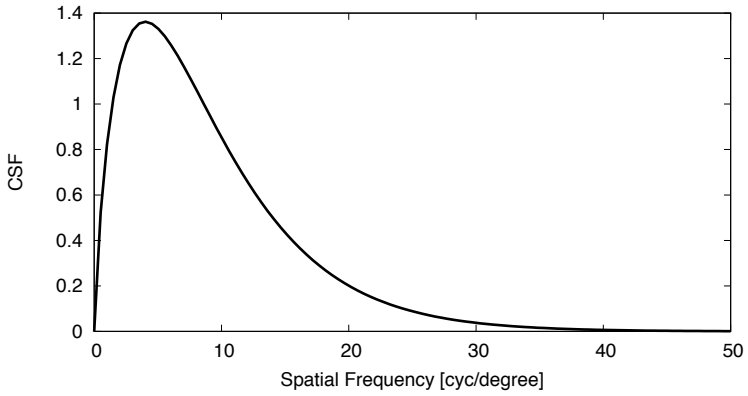


Figure 7.7: The contrast sensitivity function used in the acutance equation. The peak sensitivity is at roughly 4 cycles/degree.

modeled as

$$CSF(f) = f^{0.8} \cdot e^{-0.2f} \tag{7.8}$$

where f is in units of cycles per degree. The acutance is then given by

$$A = \frac{\iint MTF(\rho, \phi) \cdot CSF(\rho)\ d\rho\ d\phi}{\iint CSF(\rho)\ d\rho\ d\phi} \tag{7.9}$$

While acutance is a good metric for conventional photographs, it was not designed with sparse aperture systems in mind. So, the metric will be adapted for use with sparse apertures here as a demonstration, while conceding that further work is needed in sparse aperture image quality metrics. One of the main differences between sparse aperture

imaging and conventional imaging is the heavy reliance on post-processing. As such, once one frequency has enough resolution to be restored, it is better to cover more frequencies than to add more MTF at that single frequency. Thus, to account for the effect of post-processing, a similar idea from Equation 7.6 will be used.

$$f_{\text{Acutance}} = \frac{\iint \min(MTF(\rho,\phi), T) \cdot CSF(\rho)\ d\rho\ d\phi}{\iint CSF(\rho)\ d\rho\ d\phi} \tag{7.10}$$

where T is an MTF threshold indicating the level at which a frequency can be restored. The effects of T are illustrated in Figure 7.8. This threshold in reality will vary with signal-to-noise ratio, which is also a function of spatial frequency. For the sake of demonstration, it will be held to a constant value. However, when designing an actual system, it may be desirable to model the frequency content of the expected imagery to determine how SNR varies as a function of spatial frequency and vary T accordingly.

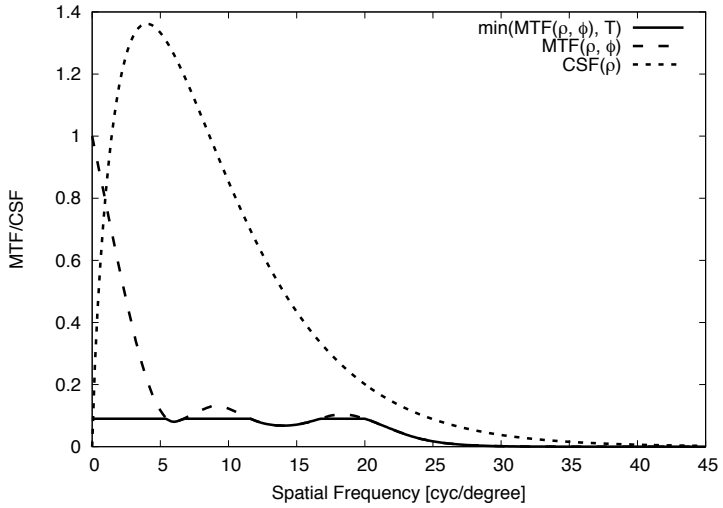


Figure 7.8: Illustration of the the effects of the threshold T in Equation 7.10, set to a value of 0.095 here.

7.4.3 Search Strategy

Once the parameterization and the fitness function have been specified, the last component of the genetic algorithm is the search strategy, or the mechanisms by which the algorithm explores the search space. In genetic algorithms, this consists of three operations: initial-

ization of the population, crossover and mutation.

7.4.3.1 Initialization

In other optimization algorithms, initialization is critical to avoiding local extrema in the fitness or cost function. It is generally less important in genetic algorithms, but must still be done properly to allow the genetic algorithm to operate. In this work, two separate search strategies were used: global search and local refinement. Global search was used to find a low-precision approximation of the optimal aperture, while local refinement was used to fine tune that approximation to obtain a high-precision optimal aperture.

The implementation of the initialization was very different in these different search strategies. In global search, the role of initialization is very minimal. The main goal is to provide an initial set of valid apertures that the crossover and mutation operators can use to construct new apertures. Simple random assignment of values to each parameter until a valid aperture is constructed is a sufficient strategy. In local refinement, the population is simply initialized as copies of the initial guess and space exploration is left to the mutation operator.

7.4.3.2 Crossover

Crossover is an operator that produces a new population member from two existing members, which have been randomly selected, according to their fitness. This operator maps well to sparse aperture design. Since MTF peaks come from the baselines between subapertures, mixing groups of apertures from previous designs should give new combinations of MTF peaks. In this work, crossover is only used in global search, as it leads to large variations in aperture design. In local refinement, the crossover operator randomly reproduces one of the input apertures. The crossover operator needs to be defined for each parameterization that is used.

As is shown in Figure 7.9, the implementation of the crossover operator is very straightforward for the (x_i, y_i, r_i) parameterization. The operator maps less well onto discrete arm parameterization, but involves randomly selecting arms from each input and then placing apertures onto each new arm. In both cases, care must be taken to maintain the fill factor through the operation. A reasonable approach is to create a "subaperture budget" for the

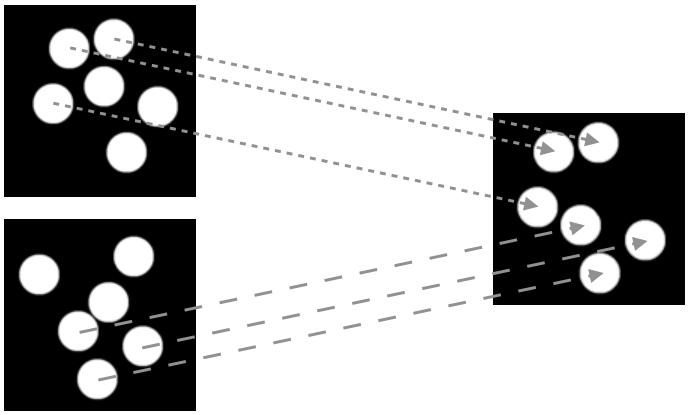


Figure 7.9: An example of the crossover operator. In the (x_i, y_i, r_i) parameterization, subapertures are randomly selected from each input aperture to construct the output aperture.

optimization. That is, the r-dimension is quantized to a small set of possible radii and each radius is given a count. Thus, the random selection can be confined to apertures of a given radius, maintaining the fill factor.

7.4.3.3 Mutation

The mutation operator functions on the output of the crossover operator and introduces random variation, thus allowing greater exploration of the search space. This has the effect of randomly changing the baselines between subapertures and thus randomly moving a subset of the MTF peaks. Mutation is used in both local and global search.

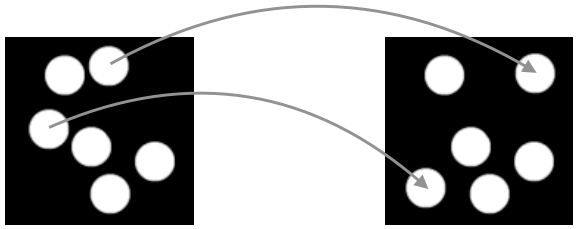
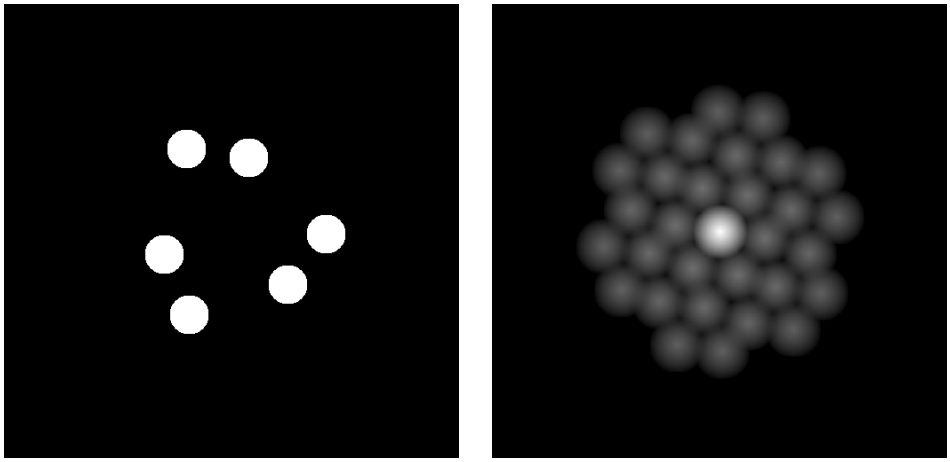


Figure 7.10: An example of the mutation operator during global search. Apertures are randomly displaced to reach into new portions of the search space.

As shown in Figure 7.10, the mutation operator on the (x_i, y_i, r_i) parameterization during global search will randomly relocate a subset of the subapertures to a new location. It will also randomly swap the radii of some apertures. In the discrete arm parameterization, it will randomly move subapertures to different arms, change their offset, move entire arms or swap radii between subapertures.

Mutation is also used as the primary driver of local refinement. Instead of randomly relocating apertures in the (x_i, y_i, r_i) parameterization, a Gaussian offset is applied to the subaperture location, while radius changes are disabled. Successive runs with standard deviations of 5%, 1% and 0.1% of the encircled diameter are used as progressive refinement of the aperture. For the discrete arm parameterization, only aperture offsets and arm angles are mutated, in a similar progressive approach.

7.5 Results



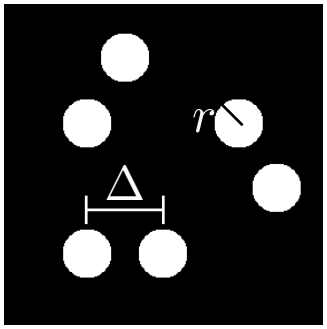
(a) Optimized aperture mask (b) Optimized MTF

Figure 7.11: Results for Golay optimization using Equation 7.3 as a fitness function ($T = 3\%$, $\gamma = 3$). The fitness function is invariant to rotation and reflection transformations.

7.5.1 Golay Validation

The first set of results use the (x_i, y_i, r_i) parameterization to replicate the results of [Golay, 1971] and [Miller et al., 2007]. In order to compare directly to the previous work, all of the subaperture radii, r_i, were set to a constant value, producing a minimum fill factor of 18%. This value was taken from [Fiete et al., 2002], as it was in the middle of the range of fill factors that produced acceptable image quality for the Golay-6 design. As can be seen from Figure 7.11, the aperture and its MTF match the Golay-6 design, which was shown in Figure 7.5 (b). This result was produced with the fitness function that was given in Equation 7.3. Using Equation 7.6 with an exponential decay weighting function will give nearly identical results.

Shown in Figure 7.11, the result obtained from the genetic apertures approach is a rotated version of the Golay-6 aperture and MTF. However, upon inspection of both Equation 7.3 and Equation 7.6, it can be seen that the optimization is invariant to both rotation and reflection. Equation 7.3, the direct approach, depends on the number of resolved frequencies, which is constant under any rotation or reflection transforms, and the distances between the subapertures, which are also maintained under these transforms. Equation 7.6, the weighting function approach, has a weighting function that only depends on radial frequency, ρ, thus giving the invariance to rotation and reflection. So, a rotated result, such as the one seen in Figure 7.11, is an equally valid aperture.



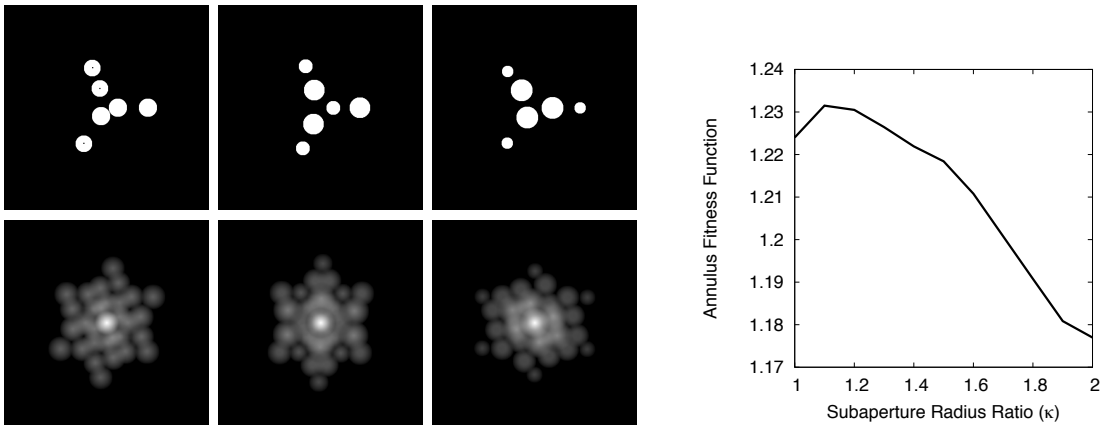
$$s = \frac{\Delta}{2r}$$

Figure 7.12: The expansion factor of a Golay aperture is defined as the ratio between the spacing between pairs of closest subapertures and the diameter of each subaperture.

Building on Golay's work, Miller et. al. in [Miller et al., 2007], examined the effect of

the expansion factor, s, illustrated in Figure 7.12. The expansion factor is the ratio of the spacing between the pairs of closest subapertures and the diameter of each subaperture. So, the minimum case, where two subapertures touch, is $s = 1$ and spacing increases as s increases. In [Miller et al., 2007], expansion factors were examined in increments of 0.1 and it was found that voids in the MTF occurred at values of s over 1.6. A void was declared to be when the MTF fell below some threshold value, which was set to 3%. This is the same parameter as the MTF threshold, T in Equation 7.3. To match the result in [Miller et al., 2007], T was set to 3% in the genetic apertures approach. This produced the result in Figure 7.11 and the expansion factor was computed to be 1.62, matching the result from [Miller et al., 2007]. If the minimum MTF value, T, is increased, it is expected that the aperture spacing will need to decrease, so that the valleys between peaks will not fall below T. This expectation matches observed behavior. If T is set to 5%, the resulting optimal expansion factor will be 1.54 and if it is set to 7%, an expansion factor of 1.46 is obtained.

7.5.2 Discrete Annulus



(a) $\kappa = 1.1$ (b) $\kappa = 1.5$ (c) $\kappa = 1.9$ (d) f_{Annulus} as a function of κ

Figure 7.13: Sample results for the discrete annulus optimization. The three shown results are representative of the different designs that result from varying the subaperture radius ratio, κ. Each design features a roughly isotropic ring around the central peak. The aperture with $\kappa = 1.1$ best optimized the fitness function. Central dots indicate the smaller subapertures in the $\kappa = 1.1$ result.

Given that the genetic apertures approach can reproduce a known result, it can be applied to new design problems with the expectation of valid results. This set of results uses the discrete arm parameterization to maximize the annulus fitness function given in Equation 7.7. For these results, three arms were used, as this is likely the most practical for construction and previous studies have examined the Tri-arm 9 design [Fiete et al., 2002]. While existing work on the Golay apertures used subapertures of a constant size, there is no reason to maintain that constraint here. In fact, having apertures of unequal size will result in MTF peaks of a constant magnitude over a small range of frequencies, which might be desirable for reproducing the annulus's constant MTF properties. For these results, the six subapertures were divided into two groups of three, with two different radii, and a constant fill factor was maintained. The radii were computed as

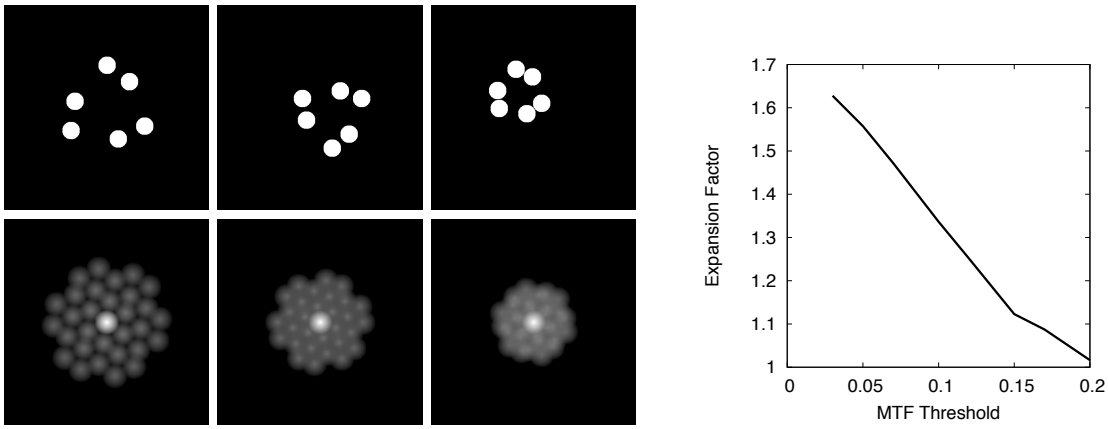
$$r_1 = r_{\text{enc}} \cdot \sqrt{\frac{F}{3(\kappa^2+1)}} \tag{7.11}$$

$$r_2 = \kappa \cdot r_1 \tag{7.12}$$

where r_{enc} is the radius of the circle enclosing the subapertures, F is the fill factor and κ is the ratio of the radii of the two sets of subapertures.

A subset of the results obtained with the annulus fitness function are given in Figure 7.13. At first inspection, the MTFs might have little resemblance to the annulus MTF. This is to be expected, as there were many restrictions placed upon these apertures, namely that they must be composed of six discrete circular subapertures and must be located on three arms intersecting at a common point, a far cry from the original annulus design. As such, it is not reasonable to expect a close approximation of the annulus MTF. Upon careful inspection, the tradeoff made by the optimization is clear. While these MTFs are not isotropic or smooth in the periphery, each design possesses a ring of peaks around the core that is approximately both constant and isotropic. So, while these apertures could not reproduce the annulus MTF over the whole frequency range, they were able to do so at the low to low-medium frequencies, which are the most important to visual perception. Interestingly, the results for this function tended to place the larger subapertures towards the center of the array. This contrasts to the results of the Golay fitness function with this parameterization and the armed arrays in [Breckinridge et al., 2008], which placed smaller subapertures towards the center.

7.5.3 Acutance



(a) $T = 3\%$ (b) $T = 12\%$ (c) $T = 20\%$ (d) Expansion factor, s, as a function of T

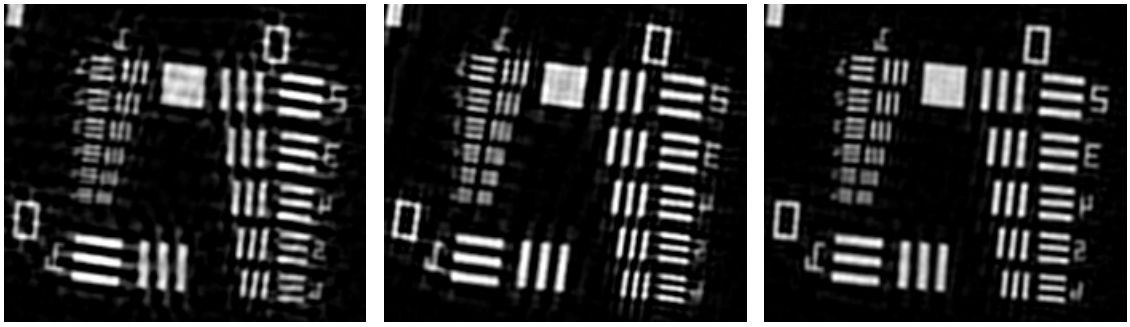
Figure 7.14: Sample results for the acutance optimization. The value of the MTF threshold, T directly influences the expansion factor, and thus fill factor, of the resulting Golay aperture.

The final set of results uses the (x_i, y_i, r_i) parameterization to maximize the acutance fitness function, given in Equation 7.10. For these results, it was assumed that the image would be viewed under normal conditions, where the Nysquist frequency of the pixel grid would match the maximum acuity of the human visual system, 60 cycles/degrees. As with the Golay fitness function, the MTF threshold, T can be used to control how much subapertures spread out over the available range. As with the Golay optimization, the minimum fill factor was set to 18%, although without a non-redundancy term, it is expected that the optimization will produce an aperture with a higher fill factor.

As can be seen in Figure 7.14, the optimal apertures for the acutance fitness function are the Golay apertures. This is not surprising, as the Golay apertures are the most efficient at covering every frequency under some cutoff frequency. The behavior as a function of the MTF threshold, T is of great interest in this optimization, however. The MTF threshold in this context refers to the minimum value of the MTF that can be restored in post-processing with some acceptable level of artifacts. An example of these artifacts was shown in Figure 7.5 (f). In this example, a Golay array with an expansion factor of

1.6 was used, equivalent to a T of 3% here. So, if that level of artifacting is deemed too high for a specific system, a higher value of T can be used to affect the expansion ratio and thus image quality. This behavior also matches the findings of [Breckinridge et al., 2008], who found that higher fill factors were needed to resolve low-contrast features, as T directly effects the MTF minimum level and thus contrast resolution. Subapertures of different radii were also examined here, as in the previous section, but did not result in any gains to the acutance fitness function.

While acutance is a measure of perceived image quality, it cannot be concluded that the Golay apertures yield the optimal image quality. Acutance has been derived by the IEEE Cell Phone Image Quality Group [Baxter et al., 2012] for systems with a monolithic aperture functions. There is no mention of inverse filtering in this metric and it does not account for inverse filtering artifacts. As such, the metric cannot be assumed to predict sparse aperture image quality and the just-noticeable-difference scale in the acutance metric is almost certainly invalid for sparse aperture imagery. As such, more research into a proper sparse aperture image quality metric is needed before declaring the Golay apertures optimal for perceived image quality, although the authors do think it likely.



(a) Simulated restored Golay image ($s = 1.6$)

(b) Simulated restored discrete annulus image ($\kappa = 1.1$)

(c) Simulated restored acutance image ($T = 0.12$)

Figure 7.15: Images simulated with optimized apertures from three fitness function. Each aperture was given the same amount of defocus (-0.5 waves), which was not corrected for in the inverse filter. The bandpass ranged from 0.5 to 1 microns. (a) and (c) used the (x_i, y_i, r_i) parameterization, while (b) used the discrete arm parameterization. All three optimizations used the same amount of glass.

7.5.4 Simulated Image Comparison

Given the results obtained in the previous sections, images can be simulated using derived optimal apertures. This simulation was done with the modeling approach in [Salvaggio et al., 2015]. If the optimizations were performed correctly, the effects of the different fitness functions should be seen in the output images. The goal of the annulus fitness function was to give a broad frequency support, while also minimizing post-processing artifacts. The goal of the acutance fitness function was to maximize perceived image quality, and thus the resulting image should subjectively look the best of the apertures to an observer. The Golay fitness function should result in the most resolution out of all of the designs, but since it does not account for post-processing, the resulting image should also have the most artifacting.

Post-processing was done using a Weiner filter of the form

$$W(\xi,\eta) = \frac{H^*(\xi,\eta)}{|H(\xi,\eta)|^2 + \alpha \cdot |S(\xi,\eta)|} \tag{7.13}$$

where H is the system OTF, α is a weighting term and S is the Fourier transform of a Laplacian convolution kernel and acts as a smoothness term. α was set to 1×10^{-3} for each image, so that all images were restored comparably. This value of α is an aggressive restoration to be used in low-noise scenarios. Since the restoration filter accounts for wavefront error, artifacting will be minimal with perfect knowledge of the wavefront error. However, this is not a realistic scenario for space-based remote sensing systems. As such, some wavefront error will be added to the system during the imaging process, but will not be corrected for in the restoration. In these simulations, -0.5 peak waves of defocus was added to all three apertures, so as to ensure a fair comparison. This value was selected to highlight the differences in artifacting between the three designs.

As can be seen from Figure 7.15, the results generally match expectations. The Golay aperture exhibits the most spatial frequency support, as the most tri-bar targets can be resolved in Figure 7.15 (a). However, this comes at the cost of post-processing artifacts, which affect the Golay aperture more so than the other apertures. The discrete annulus aperture offers a balance between resolution, contrast and artifacts. While artifacts do exist, they have a smaller magnitude than the Golay design. The contrast of the tri-bar

targets is almost identical to the Golay design, however, one less tri-bar target can be resolved. As was seen in Figure 7.13 (a), the discrete annulus has inconsistent frequency support in the higher frequencies and the cutoff frequency is non-isotropic. This was the tradeoff for obtaining constant and isotropic frequency support in the low to low-medium frequency range. So, while the image performance seen here does not reach the level of the theoretical annulus, it showcases the tradeoffs that must be made to produce a constructible array. In addition, the layout of the discrete annulus aperture was confined to three arms. This constraint was not applied to the other two designs.

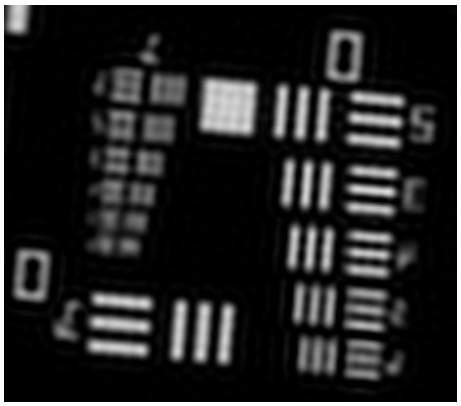


Figure 7.16: Simulated restored image from a monolithic circular aperture with the same amount of glass as the optimized apertures used in Figure 7.15. The circular aperture results in a blurrier image overall and cannot resolve the last tri-bar series.

Finally, the acutance result is the most visually pleasing of the three images after restoration, which was the intention of the optimization. The cutoff resolution is roughly equal to that of the discrete annulus aperture, but the tri-bars have less contrast and the image is smoother overall. However, the lack of any post-processing artifacts gives the acutance aperture the highest perceived image quality. In addition, this fitness function has the most intuitive tunable parameter, which can be set to eliminate post-processing for a given system's noise characteristics. Despite the differences between these sparse aperture designs, all three apertures have much higher resolution than a monolithic aperture with the equivalent amount of glass, shown in Figure 7.16, which cannot resolve any tri-bars in the last series.

7.6 Conclusions

In this chapter, an improved sparse aperture design framework was introduced, based on genetic algorithms. This framework relaxed a number of computational constraints from previous work in sparse aperture optimization. In addition, a small set of fitness functions were presented that targeted incoherent sparse aperture image quality, a problem not addressed in previous optimization studies. An initial validation study was performed by replicating the Golay-6 sparse aperture design in the new framework. This was achieved by simply replicating the design goals from Golay's work in the genetic apertures framework. When similar parameters were used to previous work, identical results were obtained.

Once the optimization framework was validated, it was applied to two design problems as demonstrations of the framework's potential use cases. The first was to replicate the attributes of another well-known sparse aperture design, the annulus. While the theoretical annulus has desirable image quality properties, it is likely impractical for construction and launch. So, a set of sparse apertures were optimized that reduced the image artifacts seen in Golay-6 imagery, while also being practically constructible with a foldable array. Such an aperture would be desirable in an application that required high frequency information, but could not accept Golay levels of artifacting. Finally, an aperture was optimized for acutance, a measure of perceived image quality. The result was a set of Golay apertures, although with higher fill factors and lower expansion factors than normally seen in sparse aperture studies. The simulated imagery subjectively looked better than the other designs, as it lacked any restoration artifacts, but came at the cost of a lower cutoff frequency. Such an aperture would be desirable if the system's main design goal was visually pleasing imagery, such as wide area coverage for a consumer geospatial information system (GIS) product.

Given this demonstration of the genetic apertures framework, the obvious future work is to find a fitness function that directly relates to image quality of restored sparse aperture imagery. Unlike sparse antenna array design, such as in [Marcano and Durán, 2000], there is not an accepted metric for the quality of a sparse aperture imaging array. Conventional image quality predictors, such as the General Image Quality Equation [Leachtenauer et al., 1997] are not applicable to sparse aperture systems. The work in [Miller et al., 2007] put forward some MTF metrics that are targeted towards sparse apertures. This work pro-

posed using existing image quality metrics, such as acutance, while bounding the MTF to account for the effect of post-processing. There has been much examination of the "effective diameter" of a sparse aperture array. Perhaps these factors can be combined to correlate with image quality for a specific application, such as perceived image quality or target detection performance.

In addition to image quality concerns, there are additional practical concerns that might need to factor into the optimization. For instance, the cost of the system will increase as a function of the complexity of the system. For monolithic systems, the cost tends to scale by at least the square of the diameter [AWMA and SPIE, 1996]. For a sparse system, additional factors, such as the alignment and mounting system for all of the mirrors, will factor into the cost. Integrating cost into the fitness function will allow optimizations that explore the trade space between cost and image quality. Besides cost, increasing the complexity of a sparse aperture design also increases the necessary engineering effort to both build the system and deploy it after launch. This research set forward the method of enforcing a practical design through the parameterization. In the case of a more flexible design than the discrete arm design presented here, a construction practicality term in the fitness function may also be useful.

While finding fitness functions targeted towards sparse aperture image quality, cost or construction practicality is obvious future work, there are still technical improvements that can be made to the genetic apertures framework. The most obvious issue is spectral performance. All optimizations done in this work were performed assuming monochromatic illumination. Increasing the bandpass has the effect of radially blurring the MTF. While this will not have a large effect on the optimized apertures, it will likely allow the apertures to spread slightly further apart, as the voids will be filled due to the spectral blurring. Using the monochromatic approximation will result in a conservative design. While the framework is perfectly capable of doing spectral simulations currently, an efficient approximation method is needed, so as to beat the linear runtime tradeoff with spectral resolution. Robustness to wavefront aberration, such as piston/tip/tilt or coma, is another interesting direction for future work, as such aberration would surely exist in a real sparse aperture system. In addition, all current optimizations have a constant number of subapertures, which was set to six in this work. It would be desirable for the framework

to be able to vary the number of subapertures, so that the user does not need to run multiple optimizations with different numbers of sparse apertures and manually compare the results. For a space-based system, only a small number of subapertures is likely practical for construction, so this would be a tractable problem.

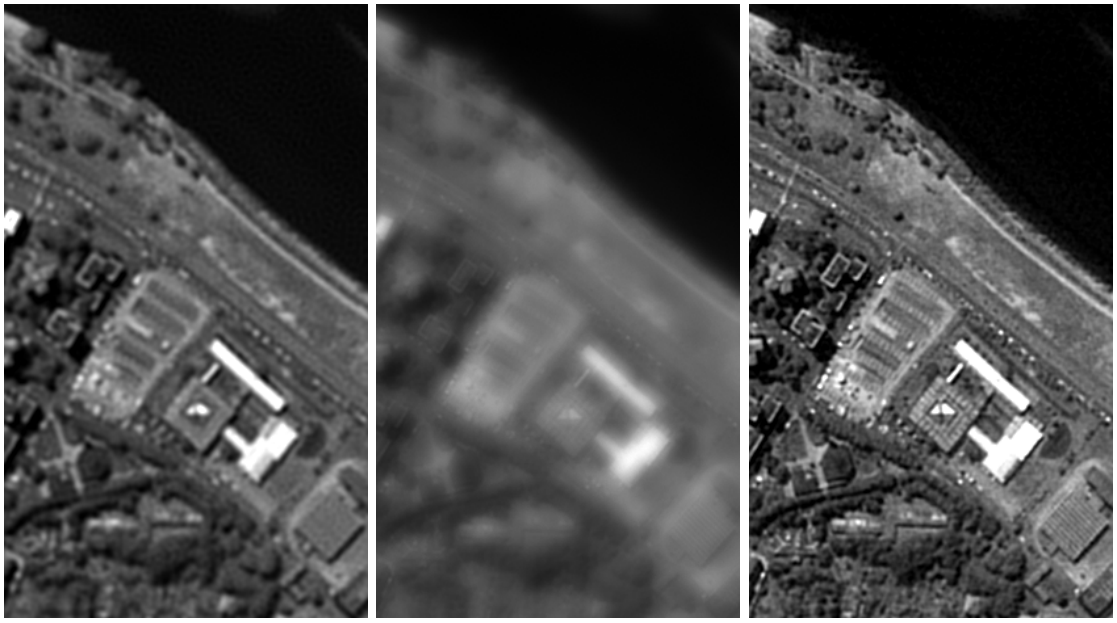
Chapter 8

Conclusions

In the previous chapters, a number of research efforts have been presented related to the challenges of image quality modeling for non-conventional aperture systems. The efforts have focused on sparse aperture systems, but the methods developed can be applied to alternative aperture designs, including systems with segmented apertures, such as the James Webb Space Telescope, or systems with phase elements inserted at the aperture plane. In this chapter, there will be a brief summary of the conclusions reached in each research effort, along with a broader discussion of the impact of these efforts and suggestions for future work.

8.1 Modeling Approach

The sparse aperture image quality model was the core of this research and its implementation was described in Chapter 4. An overview of the model was given in Figures 4.1 and 4.2. This model was primarily based off of the initial work of [Introne et al., 2005], but was built from scratch for this research with a couple of important improvements. First among these improvements was the ability to model off-axis aberrations, such as coma, which vary over the field of view of the system. This was a necessary improvement due to the coma inherent in the laboratory system used for later experiments. Improvements were also made to the aperture specification system, which allowed the model to accept a wider range of apertures outside of existing well-known designs. This improvement allowed for the optimization study in Chapter 7.



(a) Restored image with circular aperture (b) Unprocessed image with sparse aperture (c) Restored image with sparse aperture

Figure 8.1: Pavia hyperspectral data set as would be seen through a sparse aperture telescope compared to a circular aperture telescope with the same amount of glass.

While the vast majority of modeled images shown in this research have been of binary patterns, such as edge targets of USAF-1951 resolution targets, the sparse aperture model is also capable of accepting hyperspectral images, either real or simulated, as input radiance data. Figure 8.1, shows degraded versions of a portion of the common Pavia hyperspectral data set. In this figure, versions are shown through two systems using the same amount of glass, one circular and one Golay-6 sparse aperture. When looking at restored images from both cases, the resolution gains attainable with a sparse aperture are clear, manifesting as higher contrast, sharper features and increased detail on building roofs, parking lots and trees. The ability to simulate imagery from these systems will be critical to future studies utilizing this model.

8.2 OTF Validation

The first main contribution of this research is the validation of the modeling approach, which begins with the study described in Chapter 5. In that chapter, an experiment to validate the OTF predictions of the model is described, along with the laboratory setup to facilitate that experiment. This setup involves an off-axis parabolic collimator and a small-scale imaging system, which utilizes masks to create the sparse aperture effect. The use of masks was inspired by [Zhou et al., 2009], but the laboratory system used here offers a number of advantages. These advantages are the use of broadband illumination, the ability to introduce extended scenes and the ability to introduce controllable wavefront error. Previous sparse aperture validation studies had only used monochromatic illumination when measuring MTFs, however, larger bandpasses are useful in sparse aperture design in order to offset the low signal-to-noise performance due to fill factor. In order to model a realistic imaging scenario, the effects of wavelength on the system's OTF need to be modeled and thus, the ability to use broadband illumination in the validation study was a necessary improvement. In addition, the ability to introduce extended scenes aided in both this experiment, by facilitating slanted-edge measurements, and more importantly, the artifact validation experiment.

The results of this experiment were shown in Figure 5.24. These results showed that the MTFs measured in the laboratory matched the predicted MTFs obtained from the model. This experiment was repeated three times for three different levels of wavefront error and the results matched the model in each case. Sources of error were small in these experiments, but were explainable through a number of factors, such as uncertainty in the wavefront error of the laboratory system, uncertainty in the spectral transmission of each element in the system, MTF of the back-end imaging system, and image noise propagating through MTF measurement techniques. Even with all of these sources of error, RMS error between the measured and modeled MTFs were all consistently around or below 1% MTF. When the measured and modeled MTFs were used to restore slanted edge imagery, it was shown that the relative edge responses were equivalent to within the limits of human perception. Given these results, it was concluded that the model was correctly modeling the sparse aperture OTF, complete with the effects of spectral bandpass and wavefront error. This work was published and presented as [Salvaggio et al., 2015].

8.3 Artifact Validation

Once the basic OTF predictions of the model were validated, validation of post-processing artifacts was attempted. While post-processing artifacts have been predicted in optical sparse aperture imagery before, such as in [Block, 2005], there has not been a study that tries to compare the predicted artifacts to those seen in real data, which is a necessary validation that must be performed before trusting model predictions for use in trade studies. As such, Chapter 6 describes an experiment to produce artifacting in real sparse aperture imagery and compares this artifacting to model predictions. In addition, a very simple trade study was presented, addressing spectral band selection. There were two goals in the analysis of both of these studies. The first was to determine whether the model was predicting the exact artifacts seen in real data and the second was to determine for a given set of parameters, if the model was producing the right level of artifacting. That is, the model should be producing the right degradation in image quality, even if the produced artifacts do not match exactly.

Two experiments were described in Chapter 6, corresponding to the two major drivers of artifacting in sparse aperture imagery: uncertainty in wavefront error and uncertainty in the system's spectral weighting function. To a reasonable confidence, the laboratory system's spectral weighting function could be measured with a spectrometer and monochrometer. As such, it was possible to set up an experiment where the true spectral weighting function was used in restoration, while uncertainty in the wavefront error was introduced. Modeled and measured images were produced of a USAF-1951 target and 1D profiles were extracted over each tri-bar target. Profiles were compared on a point-to-point basis to test for exact matching. Error was shown to be approximately 5% of the peak intensity. Inspection of the results showed that while measured and modeled artifacts were close, they were not a perfect match. However, artifact area and peak height were also measured in both data sets and showed good agreement. As such, it is likely that the image quality impact of these artifacts were correctly predicted by the model, although such a conclusion is not entirely certain without further research into a sparse aperture image quality metric.

While the spectral weighting function of the lab setup was measurable, the wavefront error was not, due to the limited budget available for this research. As such, a different approach was taken to validate the effects of the spectral weighting function. Two sys-

tems were simulated with the use of spectral filters. A highly spectrally-varying object was created using a dual-bandpass filter in front of the illumination source. The USAF-1951 target was then imaged under different conditions. A broadband system used the entire bandpass of the imaging detector. A two-band system was simulated by collecting two images through a shortpass and longpass filter with the same cutoff frequency. Restoration of all three images was performed without knowledge of the filtered illumination spectrum and the longpass and shortpass images were then summed to restore the signal-to-noise ratio of the broadband image. This creates a scenario of two systems with identical wavefront error, but differing amounts of error in the spectral weighting function, overcoming the limitation of not knowing the wavefront error. In both the measured and modeled case, the multi-band image showed a decrease in artifacting. More importantly, this decrease was roughly consistent between the measured and the modeled data set.

The main conclusion of this study was that the post-processing artifacts predicted by the sparse aperture model are, in fact, real and will appear in actual imagery. For a given amount of wavefront error, real artifacts will appear with approximately the same height and area as those predicted by the model, although their shape and position might vary slightly due to noise and pixel phasing effects. Unfortuantely, due to the lack of precise knowledge of the wavefront error, a perfect match between modeled and measured data was not possible. Regardless, the intuition that artifacts arose from errors in the wavefront error estimate and spectral weighting function estimate was confirmed. Finally, it was shown that careful selection of spectral bands can reduce the error due to the implicit gray-world assumption in restoration, resulting in a decrease in artifacting. By showing that this improvement in artifacting performance was accurately predicted by the model, it was shown that the model can be used in future trade studies about artifact performance with reasonable confidence.

8.4 Aperture Layout Optimization

Chapter 7 described the final study in this research and focused on optimizing the spatial layout of sparse apertures. This optimization was performed using a genetic algorithm. The main contribution of this study was the framework of the optimization, however, a number of demonstrations were also outlined in this chapter, which was published as [Sal-

vaggio et al., 2016]. While a large number of variables may be optimized, the domain of optimization in the presented demonstrations was confined to the positions and sizes of circular subapertures.

In order to run an optimization, the user has to specify three components: the parameterization, search strategy and fitness function. In Chapter 7, two parameterization and search strategy examples were provided, along with three fitness function examples. The first study that was done was aimed as a validation experiment. [Golay, 1971] had derived a set of aperture patterns that maximized both non-redundancy and compactness. These results were reproduced in the genetic algorithm framework. Subapertures were allowed to move anywhere within a given radius and the radius of each subaperture was held constant, as it was in Golay's work. A fitness function was designed to count the number of "resolved" frequencies and balanced that with the compactness of the peaks in the MTF. Golay's results were reproduced by finding the correct balance between the two terms in the fitness function. Golay had confined his result to have three-fold symmetry, however, this constraint was not placed on the optimization and occurred naturally.

Once the genetic algorithm approach was validated, it could be used on new design problems. Another fitness function was designed to reproduce the frequency domain benefits of the annulus sparse aperture design, which has been seen in previous works. However, in this study, a practical and build-able design was desired. So, it was demonstrated that the parameterization could be used to enforce design constraints, in this case, that subapertures must be located on a discrete number of "arms". In addition, subaperture size was allowed to vary in a controlled manner, although the optimization showed that the best design had roughly constant subaperture size. The results of this study demonstrated some of the benefits of the annulus design. That is, the simulated imagery showed good modulation performance with reduced artifacting. Qualitative examination of the resulting MTFs showed that the optimized apertures emulated the annulus design at low frequencies quite well, despite the considerable design constraints of six subapertures on three arms.

The final demonstration in this study aimed to maximize perceived image quality. To this end, the acutance metric from [Baxter et al., 2012] was used as a fitness function, with

some slight modifications for sparse aperture imagery. This fitness function offered an adjustable parameter, which served to control the amount of artifacting that was deemed acceptable for a given application. The results of this study were the Golay apertures, although with lower expansion factors than normally seen in sparse aperture studies. When the expansion factor is decreased, secondary peaks in the Golay aperture condense and overlap, causing the MTF to resemble the constant disc of the annulus design. As such, it made sense that this MTF would maximize the acutance fitness function. Simulated imagery from the optimized apertures looked subjectively better than imagery from the other two, reflecting the goals of the fitness function.

The three demonstrations presented in Chapter 7 showed the adaptability of the optimization framework, even with a constrained domain of optimization parameters. By changing the fitness function, the user was able to obtain three very different telescopes, one that resolved as many frequencies as possible within a certain cutoff frequency, one that balanced artifacting, resolution and constructibility, and one that maximized perceived image quality. When combined with future metrics for sparse aperture image quality, this framework could prove very useful to future designers of sparse aperture telescopes. This design optimization framework could also be easily extended to other non-conventional aperture systems that can be compactly parameterized.

8.5 Broader Impacts

The research presented here can have impacts for the field of sparse aperture imaging, as well as imaging with other non-conventional apertures. In the field of sparse aperture imaging, this works showed the validity of the model presented in Chapter 4, which is an updated and improved version of the model presented in [Introne, 2004]. The OTF predictions of the model were validated through the use of a small-scale laboratory sparse aperture system. This validation study was more extensive than previous efforts in that it introduced both broadband illumination and controllable wavefront error. Such a setup could be used to perform broadband MTF studies for other types of non-conventional systems and the ability to introduce extended scenes allows this laboratory setup to be used for other types of calibration and testing.

Beyond simple OTF predictions, the model is capable of performing image simulations. The most unique feature of simulated sparse aperture images is ringing artifacts after post-processing. The occurrence of these artifacts in real data was also validated through a laboratory experiment. It was shown that the model can be trusted to predict the correct level of artifacting and can be used as a stand-in for real imagery in a trade study focused on these artifacts. This property can result in large cost reduction to anyone designing a sparse aperture telescope, as it can potentially eliminate a good deal of physical prototyping. Finally, given that trust in the model was established with the validation studies, it was shown that the user could design new sparse aperture systems by using model predictions in conjunction with an optimization framework, based on genetic algorithms. With further computational improvements, it could be possible to integrate artifact predictions directly into the optimization.

While all of these contributions are useful to the field of sparse aperture imaging, they are extensible to other forms of non-conventional imaging. The two most notable examples are segmented aperture imaging and phase elements. It was shown that the precursor to the model used here could be applied to segmented apertures, such as the James Webb Space telescope, by [Zelinski, 2009]. This model could be used to perform trade studies on such systems, such as spectral band selection. If a larger sparse segmented aperture were desired, this model could be used to determine which sub-mirrors could be omitted to minimize the effect on image quality. Phase elements are an even more natural extension. Phase elements can be used to alter the point spread function of a system, as in [Ruane et al., 2015]. Almost every step of the methodology in this research could be replicated by simply replacing the sparse aperture mask with a phase element. If the phase function of the element is parametrizable, it could also be optimized with the framework set forward in Chapter 7.

While the modeling framework presented and validated in this work can be critical to performing trade studies on non-conventional aperture systems, there are some limitations that need to be addressed. This model only covers a portion of the imaging chain. That is, it ignores all of the light-matter interactions that occur prior to the radiance reaching the entrance pupil of the system. Similarly, the model ignores the last steps of the imaging system, those of display, perception or analysis. As such, the model only functions as part

of an image simulation pipeline, but can be integrated with other components, such as the Digital Imaging and Remote Sensing Image Generation (DIRSIG) model, to model the entire imaging chain. However, the final quality of the simulated system will be dependent on the tools chosen to address the remainder of the imaging chain. In addition, the model presented does not explicitly model all of the necessary optics for a real sparse aperture system. In reality, a significant amount of back-end optics, such as fold mirrors and beam collectors, will be necessary. Misalignments or imperfections in these optics will cause wavefront error, such as piston, tip and tilt, among others. Once computed, these errors can be integrated into this model, however, the computation of these aberrations was not covered here. Beyond these limitations, there are numerous improvements and areas for future work.

8.6 Future Work

There are a large number of possible areas for future work based off of this research. The most obvious is the area of sparse aperture image quality metrics. While the NIIRS scale is still applicable to sparse aperture imagery, the GIQE regression fit is not. A sparse aperture image most likely should be characterized using different or additional terms for both resolution and post-processing artifacts and both the coefficient values and functional relationships are likely to change. A future study that combines predictions from this model and the methods from [Garma, 2015] would be an interesting continuation of this work. Findings from this study could then be integrated with the optimization approach from [Salvaggio et al., 2016] to produce useful new sparse aperture designs.

Future work on sparse aperture modeling is also a possibility. A key assumption in the modeling approach of this work was that the entrance pupil-reaching radiance field could be described as a two-dimensional quantity. This is a valid assumption for a satellite-sized telescope. However, if such a telescope is spread over a large enough area, then this assumption breaks down and perspective differences begin to emerge between the telescopes, resulting in a depth of field/focus for the telescope. The integration of this effect into the model would widen the potential use cases to even more exotic system designs. In Chapter 6, a number of limitations of the model and the laboratory setup became clear. The issue of pixel phasing became clear when evaluating the spectral data and the ability to

provide multi-scale input imagery to the model would be useful in examining how artifacts varied as a result of pixel phasing. The uncertainty in wavefront error was the primary reason that the model could not exactly predict the observed artifacts. A system that had more controllable wavefront error, through the use of adaptive optics equipment, such as a wavefront sensor and a deformable mirror, would allow for more precise and automated comparisons between measured and modeled data. So, while defocus and coma was the only wavefront error explicitly examined in this work, the introduction of a deformable mirror could allow for automated experiments that examined the effects of other aberrations, such as local piston, tip and tilt, among others.

While artifact predictions were directly validated, it would also be advantageous to validate entire two-dimensional image simulations. As work in hyperspectral image projectors advances, such as in [Iacchetta et al., 2015], it would at some point be worth investigating whether such a system could be integrated into the laboratory system. While the post-processing artifacts are the most unconventional portion of the image simulation outputs, it would still be a worthwhile conclusion to see that image simulations are, in fact, in agreement with real data. Such a study could either perform subjective experiments with human observers or task-based automated studies to establish whether image quality was accurately modeled.

Finally, the issue of post-processing on sparse aperture imagery could be investigated further. In this research, the post-processing methodology was held constant throughout every experiment. This allowed for analysis of the underlying effects that caused artifacting, namely wavefront error and spectral weighting function uncertainty. While improving these sources of error will generally decrease artifacting, it may also be possible to decrease overall artifacting by improving the post-processing operation. This may be accomplished through a number of means, such as examining non-linear techniques or spatially varying the linear post-processing filter. It would be useful for future studies to look into whether additional information, such as the spectral content of the scene, would result in significant decreases in artifacting.

Bibliography

Noah Block. *A sensitivity study of a polychromatic sparse-aperture system.* PhD thesis, Rochester Institute of Technology, Rochester, NY, September 2005.

Philip S Salvaggio, John R Schott, and Donald M McKeown. Laboratory validation of a sparse aperture image quality model. In *SPIE Optical Engineering+ Applications*, pages 961708–961708. International Society for Optics and Photonics, 2015.

Peter B. de Selding. DigitalGlobe Awards $307M in Contracts for WorldView-3 Satellite. *Space News*, 2010.

Jon C Leachtenauer. National imagery interpretability rating scales overview and product description. In *ASPRS/ACSM Annual Convention & Exposition Technical Papers*, volume 1, pages 262–272. ASPRS/ACSM, 1996.

Jon C Leachtenauer, William Malila, John Irvine, Linda Colburn, and Nanette Salvaggio. General image-quality equation: Giqe. *Applied Optics*, 36(32):8322–8328, 1997.

J.W. Goodman. *Introduction to Fourier Optics.* MaGraw-Hill, 3rd edition, 2005.

AWMA and SPIE. *Optical remote sensing for environmental and process monitoring: proceedings of a symposium sponsored by the Air & Waste Management Association, San Francisco, CA, September 25-27, 1995.* VIP (Series). Air & Waste Management Association, 1996. URL `https://books.google.com/books?id=hFNRAAAAMAAJ`.

Rey Jan D. Garma. *Image Quality Modeling and Characterization of Nyquist Sampled Framing Systems with Operational Considerations for Remote Sensing.* PhD thesis, Rochester Institute of Technology, Rochester, NY, July 2015.

J.R. Schott. *Remote Sensing: The Image Chain Approach*. Oxford University Press, USA, 2007. ISBN 9780195178173. URL http://books.google.com/books?id=_LYBmQEACAAJ.

Jack D. Gaskill. *Linear Systems, Fourier Transforms, an Optics*. John Wiley & Sons, New York, NY, 1978.

Roger L. Easton. *Fourier Methods in Imaging*. The Wiley-IS&T Series in Imaging Science and Technology. Wiley, 2010.

A.B. Berk, L.S. Bernstein, and D.C. Robertson. Modtran: A moderate resolution model for lowtran 7. Technical Report Tech. Rep. GL-TR-89-0122, Spectral Sciences Inc., Burlington, MA, 1989.

Robert D. Fiete, Theodore A. Tantalo, Jason R. Calus, and James A. Mooney. Image quality of sparse-aperture designs for remote sensing. *Optical Engineering*, 41(8):1957–1969, 2002. doi: 10.1117/1.1490555. URL http://dx.doi.org/10.1117/1.1490555.

P.J. Napier, A.R. Thompson, and R.D. Ekers. The very large array: Design and performance of a modern synthesis radio telescope. In *Proceedings of the IEEE*, volume 71, pages 1295–1320. IEEE, 1989.

SS Reddi. Constrained least-squares image restoration: an improved computational scheme. *Applied optics*, 17(15):2340–2341, 1978.

John R. Schott, Rolando V. Raqueno, and Carl Salvaggio. Incorporation of a time-dependent thermodynamic model and a radiation propagation model into infrared three-dimensional synthetic image generation. *Optical Engineering*, 31(7):1505–1516, 1992.

AB Meinel. Aperture synthesis using independent telescopes. *Applied Optics*, 9(11):2501–2504, 1970.

Marcel J. E. Golay. Point arrays having compact, nonredundant autocorrelations. *J. Opt. Soc. Am.*, 61(2):272–273, Feb 1971. doi: 10.1364/JOSA.61.000272. URL http://www.opticsinfobase.org/abstract.cfm?URI=josa-61-2-272.

James R. Fienup. Mtf and integration time versus fill factor for sparse-aperture imaging systems. In *Proceedings of the SPIE SPIE*, volume 4091, pages 43–47, 2000. URL http://dx.doi.org/10.1117/12.405803.

James Breckinridge, Nevin Bryant, and John Lorre. Innovative pupil topographies for sparse aperture telescopes and snr. In *SPIE Astronomical Telescopes+ Instrumentation*, pages 70133E–70133E. International Society for Optics and Photonics, 2008.

Robert E Introne, Noah R Block, and John R Schott. Comparison of monochromatic and polychromatic modeling of sparse-aperture image quality. In *Aerospace Conference, 2005 IEEE*, pages 1944–1962. IEEE, 2005.

Michael E Zelinski and John R Schott. Segmented aperture space telescope modeling used for remote sensing and image utility analysis. In *SPIE Defense, Security, and Sensing*, pages 733009–733009. International Society for Optics and Photonics, 2009.

Jason Smith. *An Evaluation of multispectral earth-observing multi-aperture telescope designs for target detection and characterization.* PhD thesis, Rochester Institute of Technology, Rochester, NY, July 2012.

Soon-Jo Chung, David W. Miller, and Olivier L. de Weck. Design and implementation of sparse-aperture imaging systems. In Howard A. MacEwen, editor, *Proceedings of the SPIE*, volume 4849 of *Highly Innovative Space Telescope Concepts*. SPIE, December 2002.

Z. Zhou, D. Wang, J. Zhao, Y. Wan, Z. Jiang, and S. Tao. Comparison of optical sparse aperture image restoration with experimental psf and designed psf. In S. Ryan, editor, *Proceedings of the Advanced Maui Optical and Space Surveillance Technologies Conference*, page E94, Wailea, Maui, Hawaii, September 2009. The Maui Economic Development Board.

Robert E. Introne. *Enhanced spectral modeling of sparse aperture imaging systems.* PhD thesis, Rochester Institute of Technology, Rochester, NY, 2004.

Robert D Fiete. Image quality and λfn/p for remote sensing systems. *Optical Engineering*, 38(7):1229–1240, 1999.

Nicholas J. Miller, Matthew P. Dierking, and Bradley D. Duncan. Optical sparse aperture imaging. *Appl. Opt.*, 46(23):5933–5943, Aug 2007. doi: 10.1364/AO.46.005933. URL `http://ao.osa.org/abstract.cfm?URI=ao-46-23-5933`.

James R. Fienup, Douglas K. Griffith, L. Harrington, A. M. Kowalczyk, and Jason J. Miller. Comparison of reconstruction algorithms for images from sparse-aperture systems. In Philip J. Bones, Michael A. Fiddy, and Rick P. Millane, editors, *Proceedings of the SPIE*, volume 4792 of *Image Reconstruction from Incomplete Data II*, 2002.

Peter D. Burns. Slanted-edge mtf for digital camera and scanner analysis. In *Proceeding of IS&T 2000 PICS Conference*, pages 135–138. IS&T, 2000.

Martin A. Fischler and Robert C. Bolles. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM*, 24(6):381–395, June 1981. ISSN 0001-0782. doi: 10.1145/358669.358692. URL `http://doi.acm.org/10.1145/358669.358692`.

Aden B Meinel, Marjorie P Meinel, and Neville J Woolf. Multiple aperture telescope diffraction images. *Applied optics and optical engineering.*, 9:149–201, 1983.

Diógenes Marcano and Filinto Durán. Synthesis of antenna arrays using genetic algorithms. *Antennas and Propagation Magazine, IEEE*, 42(3):12–20, 2000.

Kesong Chen, Zishu He, and Chunlin Han. Design of 2-dimension sparse arrays using an improved genetic algorithm. In *Sensor Array and Multichannel Processing, 2006. Fourth IEEE Workshop on*, pages 209–213. IEEE, 2006.

Kesong Chen, Xiaohua Yun, Zishu He, and Chunlin Han. Synthesis of sparse planar arrays using modified real genetic algorithm. *Antennas and Propagation, IEEE Transactions on*, 55(4):1067–1073, 2007.

Munish Rattan, Manjeet Singh Patterh, and BS Sohi. Antenna array optimization using evolutionary approaches. *Apeiron*, 15(1):78, 2008.

Philip D Henshaw and Norman R Guivens Jr. Genetic algorithms for unconventional imaging. In *SPIE's International Symposium on Optical Engineering and Photonics in Aerospace Sensing*, pages 257–265. International Society for Optics and Photonics, 1994.

Philip S Salvaggio, John R Schott, and Donald M McKeown. Genetic apertures: an improved sparse aperture design framework. *Applied Optics*, 55(12):3182–3191, 2016.

Melanie Mitchell. *An Introduction to Genetic Algorithms*. MIT Press, Cambridge, MA, USA, 1998. ISBN 0262631857.

Adam Lipowski and Dorota Lipowska. Roulette-wheel selection via stochastic acceptance. *Physica A*, 391:2193–2196, 2011.

P.E. Hart, N.J. Nilsson, and B. Raphael. A fromal basis for the heuristic determination of minimum cost paths. *IEEE Transactions on Systems Science and Cybernetics*, 4(2): 100–107, 1968.

Alan Turing. On computable numbers, with an application to the entscheidungsproblem. In *Proceedings of the London Mathematical Society*, volume 42 of *2*, pages 230–265, 1937.

Donald Baxter, Frédéric Cao, Henrik Eliasson, and Jonathan Phillips. Development of the i3a cpiq spatial metrics. In *Proc. SPIE*, volume 8293, pages 829302–829302–12, 2012. doi: 10.1117/12.905752. URL `http://dx.doi.org/10.1117/12.905752`.

Michael E. Zelinski. *A Segmented aperture space telescope modeling tool and its application to remote sensing as understood through image quality and image utility*. PhD thesis, Rochester Institute of Technology, Rochester, NY, August 2009.

Garreth J Ruane, Abbie T Watnik, and Grover A Swartzlander Jr. Reducing the risk of laser damage in a focal plane array using linear pupil-plane phase elements. *Applied optics*, 54(2):210–218, 2015.

Alexander S. Iacchetta, James R. Fienup, David T. Leisawitz, and Matthew R. Bolcar. Nonnegative matrix factorization for efficient hyperspectral image projection. In *Proc. SPIE*, volume 9611, pages 96110Y–96110Y–9, 2015. URL `http://dx.doi.org/10.1117/12.2188119`.